

Calculating Basic Statistical Procedures in SPSS: A Self-Help and Practical Guide to Preparing Theses, Dissertations, and Manuscripts

By:

John R. Slate

Ana Rojas-LeBouef

Calculating Basic Statistical Procedures in SPSS: A Self-Help and Practical Guide to Preparing Theses, Dissertations, and Manuscripts

By:

John R. Slate
Ana Rojas-LeBouef

Online:

< <http://cnx.org/content/col11292/1.6/> >

C O N N E X I O N S

Rice University, Houston, Texas

This selection and arrangement of content as a collection is copyrighted by John R. Slate, Ana Rojas-LeBouef. It is licensed under the Creative Commons Attribution 3.0 license (<http://creativecommons.org/licenses/by/3.0/>).

Collection structure revised: April 28, 2011

PDF generated: October 29, 2012

For copyright and attribution information for the modules contained in this collection, see p. 153.

Table of Contents

1 Introduction: Why a Book on Statistical Help for Graduate Students and Faculty?	1
2 Calculating Descriptive Statistics	5
3 Calculating a Nonparametric Pearson Chi-Square	25
4 Calculating Correlations: Parametric and Non Parametric	35
5 Conducting a Parametric Independent Samples t-test	55
6 Conducting a Parametric Dependent Samples t-test (Paired Samples t-test)	71
7 Conducting a Nonparametric Independent Samples t-test	85
8 Conducting a Nonparametric Paired Samples t-test	101
9 Conducting a Parametric One-Way Analysis of Variance	113
10 Conducting a Nonparametric One-Way Analysis of Variance	131
11 Standardized Skewness and Standardized Kurtosis Coefficient Calculator	149
12 Resources: Calculating Basic Statistics in SPSS	151
Attributions	153

Chapter 1

Introduction: Why a Book on Statistical Help for Graduate Students and Faculty?



NOTE: This Chapter has been peer-reviewed, accepted, and endorsed by the National Council of Professors of Educational Administration (NCPEA) as a significant contribution to the scholarship and practice of education administration. Formatted and edited in Connexions by Theodore Creighton and Brad Bizzell, Virginia Tech, Janet Tareilo, Stephen F. Austin State University, and Thomas Kersten, Roosevelt University.

This chapter is part of a larger Collection (Book) and is available at: *Calculating Basic Statistical Procedures in SPSS: A Self-Help and Practical Guide to Preparing Theses, Dissertations, and Manuscripts*²

NOTE: Slate and LeBouef have written a "companion book" which is available at: *Preparing and Presenting Your Statistical Findings: Model Write Ups*³

Authors Information

John R. Slate is a Professor at Sam Houston State University where he teaches Basic and Advanced Statistics courses, as well as professional writing, to doctoral students in Educational Leadership and Counseling. His research interests lie in the use of educational databases, both state and national, to reform school practices. To date, he has chaired and/or served over 100 doctoral student dissertation committees. Recently, Dr. Slate created a website, Writing and Statistical Help⁴ to assist students and faculty with both statistical assistance and in editing/writing their dissertations/theses and manuscripts.

¹This content is available online at <<http://cnx.org/content/m37280/1.2/>>.

²<http://my.qoop.com/store/NCPEA-Publications-1781472103076212/>

³<http://my.qoop.com/store/NCPEA-Publications-1781472103076212/>

⁴<http://cnx.org/content/m37280/latest/www.writingandstatisticalhelp.com>

Ana Rojas-LeBouef is a Literacy Specialist at the Reading Center at Sam Houston State University where she teaches developmental reading courses. She recently completed her doctoral degree in Reading, where she conducted a 16-year analysis of Texas statewide data regarding the achievement gap. Her research interests lie in examining the inequities in achievement among ethnic groups. Dr. Rojas-LeBouef also assists students and faculty in their writing and statistical needs on the Writing and Statistical website, Writing and Statistical Help⁵

Editors Information

Theodore B. Creighton, is a Professor at Virginia Tech and the Publications Director for NCPEA Publications⁶, the Founding Editor of Education Leadership Review,⁷ and the Senior Editor of the NCPEA Connexions Project.

Brad E. Bizzell, is a recent graduate of the Virginia Tech Doctoral Program in Educational Leadership and Policy Studies, and is a School Improvement Coordinator for the Virginia Tech Training and Technical Assistance Center. In addition, Dr. Bizzell serves as an Assistant Editor of the NCPEA Connexions Project in charge of technical formatting and design.

Janet Tareilo, is a Professor at Stephen F. Austin State University and serves as the Assistant Director of NCPEA Publications. Dr. Tareilo also serves as an Assistant Editor of the NCPEA Connexions Project and as a editor and reviewer for several national and international journals in educational leadership.

Thomas Kersten is a Professor at Roosevelt University in Chicago. Dr. Kersten is widely published and an experienced editor and is the author of Taking the Mystery Out of Illinois School Finance⁸, a Connexions Print on Demand publication. He is also serving as Editor in Residence for this book by Slate and LeBouef.

1.1 Introduction: Why a Book for Helping Students and Faculty with SPSS and Writing Help?

In the past two decades of teaching basic and advanced statistical procedures, we have observed student after student who experienced difficulty with using the Statistical Package for the Social Sciences (SPSS) and with interpreting the voluminous output generated by SPSS. These difficulties, along with statistics anxiety experienced by many students, led us to develop a specific and detailed set of steps for students to follow. Students reported to us, over and over, how helpful the point-and-click steps were to them in allowing them to use SPSS. Some students, even with the steps, still managed to experience difficulty in being able to use SPSS successfully. As a result, we generated screenshots for every major point-and-click step. This combination of steps and screenshots has met with excellent student satisfaction and, most importantly for us as instructors, has enhanced their ability to be successful in using SPSS.

We have written this textbook in hopes of facilitating individuals' success in using SPSS for their statistical analyses and in interpreting the SPSS output properly. Graduate and undergraduate students who take a statistics course in which SPSS is used will find these steps and screenshots to be very practical and very easy to follow. Doctoral students, who completed their statistics course years ago, but who are now working on their dissertation data analysis will find this textbook to be a practical step-by-guide. Finally, faculty members who engage in scholarly activities but are years removed from their own statistics courses will find this textbook to be helpful.

We hope that you find our materials helpful to you in your use of SPSS and in your interpretation of SPSS output. This textbook reflects our efforts and interests in making statistical analysis less threatening and less anxiety-producing than many persons find it to be. Currently, great emphasis is placed on accountability in

⁵<http://cnx.org/content/m37280/latest/www.writingandstatisticalhelp.com>

⁶<http://www.ncpeapublications.org>

⁷<http://ncpeapublications.org/about-elr.html>

⁸<http://cnx.org/content/col10606/latest/>

educational settings. Being able to analyze data, of which an abundance clearly exists, in an interpretable way is essential, especially if we want to make the educational lives of our students better.

John R. Slate, Sam Houston State University

Ana Rojas-LeBouef, Sam Houston State University

*CHAPTER 1. INTRODUCTION: WHY A BOOK ON STATISTICAL HELP
FOR GRADUATE STUDENTS AND FACULTY?*

Chapter 2

Calculating Descriptive Statistics¹



NOTE: This Chapter has been peer-reviewed, accepted, and endorsed by the National Council of Professors of Educational Administration (NCPEA) as a significant contribution to the scholarship and practice of education administration. Formatted and edited in Connexions by Theodore Creighton and Brad Bizzell, Virginia Tech, Janet Tareilo, Stephen F. Austin State University, and Thomas Kersten, Roosevelt University.

2.1

This chapter is part of a larger Collection (Book) and is available at: *Calculating Basic Statistical Procedures in SPSS: A Self-Help and Practical Guide to Preparing Theses, Dissertations, and Manuscripts* ²

NOTE: Slate and LeBouef have written a "companion book" which is available at: *Preparing and Presenting Your Statistical Findings: Model Write Ups* ³

Authors Information

John R. Slate is a Professor at Sam Houston State University where he teaches Basic and Advanced Statistics courses, as well as professional writing, to doctoral students in Educational Leadership and Counseling. His research interests lie in the use of educational databases, both state and national, to reform school practices. To date, he has chaired and/or served over 100 doctoral student dissertation committees. Recently, Dr. Slate created a website, Writing and Statistical Help⁴ to assist students and faculty with both statistical assistance and in editing/writing their dissertations/theses and manuscripts.

Ana Rojas-LeBouef is a Literacy Specialist at the Reading Center at Sam Houston State University where she teaches developmental reading courses. She recently completed her doctoral degree in Reading, where she conducted a 16-year analysis of Texas statewide data regarding the achievement gap.

¹This content is available online at <<http://cnx.org/content/m37276/1.9/>>.

²<http://my.qoop.com/store/NCPEA-Publications-1781472103076212/>

³<http://my.qoop.com/store/NCPEA-Publications-1781472103076212/>

⁴<http://cnx.org/content/m37276/latest/www.writingandstatisticalhelp.com>

Her research interests lie in examining the inequities in achievement among ethnic groups. Dr. Rojas-LeBouef also assists students and faculty in their writing and statistical needs on the Writing and Statistical website, Writing and Statistical Help⁵

Editors Information

Theodore B. Creighton, is a Professor at Virginia Tech and the Publications Director for NCPEA Publications⁶, the Founding Editor of Education Leadership Review,⁷ and the Senior Editor of the NCPEA Connexions Project.

Brad E. Bizzell, is a recent graduate of the Virginia Tech Doctoral Program in Educational Leadership and Policy Studies, and is a School Improvement Coordinator for the Virginia Tech Training and Technical Assistance Center. In addition, Dr. Bizzell serves as an Assistant Editor of the NCPEA Connexions Project in charge of technical formatting and design.

Janet Tareilo, is a Professor at Stephen F. Austin State University and serves as the Assistant Director of NCPEA Publications. Dr. Tareilo also serves as an Assistant Editor of the NCPEA Connexions Project and as a editor and reviewer for several national and international journals in educational leadership.

Thomas Kersten is a Professor at Roosevelt University in Chicago. Dr. Kersten is widely published and an experienced editor and is the author of Taking the Mystery Out of Illinois School Finance⁸, a Connexions Print on Demand publication. He is also serving as Editor in Residence for this book by Slate and LeBouef.

2.2 Calculating Descriptive Statistics

In this set of steps, readers are provided with directions on calculating basic measures of central tendency (i.e., mean, median, and mode), measures of dispersion (i.e., standard deviation, variance, and range), and measures of normality (i.e., skewness and kurtosis). For detailed information regarding the advantages and limitations of each of the measures cited, readers are referred to the Hyperstats Online Statistics Textbook at <http://davidmlane.com/hyperstat/>⁹ or to the *Electronic Statistics Textbook* (2011) at <http://www.statsoft.com/textbook/>¹⁰

Step One

First check the accuracy of your dataset.

✓ Analyze

* Descriptive Statistics

* Frequencies

⁵<http://cnx.org/content/m37276/latest/www.writingandstatisticalhelp.com>

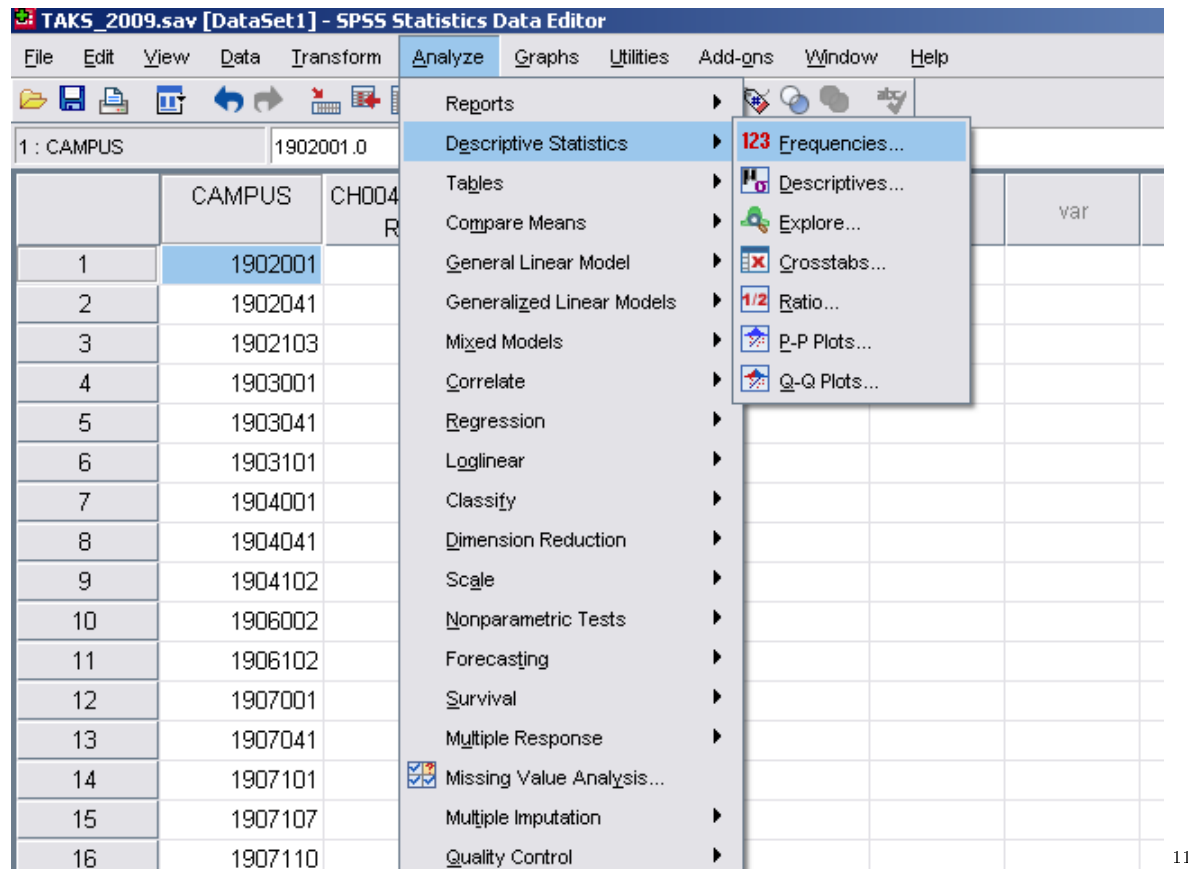
⁶<http://www.ncpeapublications.org>

⁷<http://ncpeapublications.org/about-elr.html>

⁸<http://cnx.org/content/col10606/latest/>

⁹<http://davidmlane.com/hyperstat/>

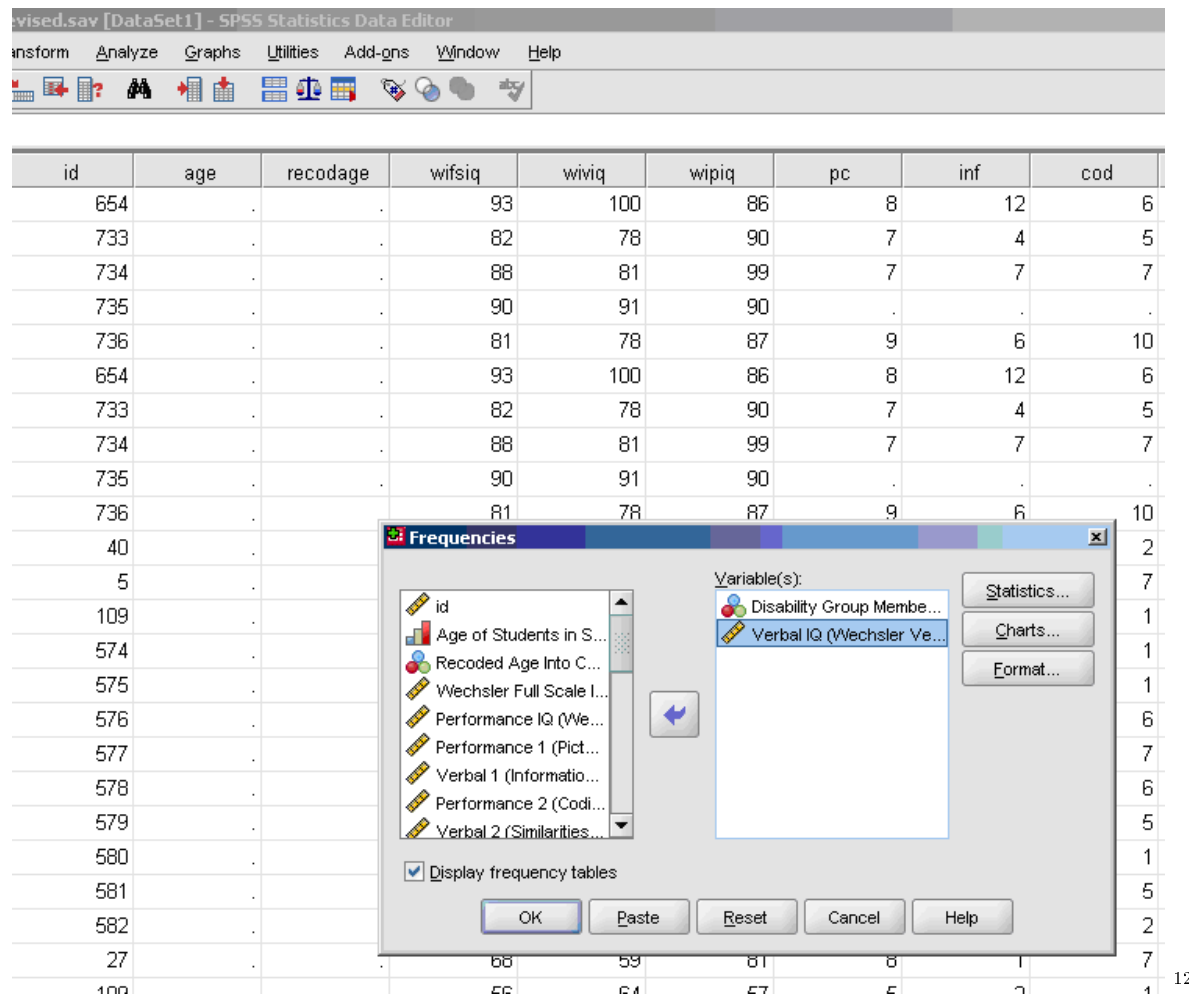
¹⁰<http://www.statsoft.com/textbook/>



11

- ✓ Move over the independent variable/s
- ✓ Move over the dependent variable/s
- ✓ OK

¹¹<http://cnx.org/content/m37276/latest/figure1.1-thumb.png/image>



* Uncheck the display "frequency tables" so that you are not provided with the frequencies of your data every time descriptive statistics are obtained.

Now check your output to see that the values for each of the variables is within the possible limits (e.g., 1 and 2 for gender). If your dataset is inaccurate, correct any inaccuracies before calculating any statistics.

To calculate descriptive statistics:

✓ Analyze

* Descriptive Statistics

* Frequencies

* Move over the dependent variable/s

* Do **NOT** move over the independent variable/s or any string variables

* Statistics

¹²<http://cnx.org/content/m37276/latest/figure1.2.PNG/image>

The screenshot displays the SPSS Statistics Data Editor window with a dataset named 'practice SPSS database revised.sav'. The dataset has 25 rows and 10 columns: group, id, age, recodage, wfsiq, wvliq, wpiq, pc, inf, and cod. A 'Frequencies' dialog box is open, showing the variable 'Verbal IQ (Wechsler Ve...' selected. The dialog also includes options for 'Statistics...', 'Charts...', 'Format...', and 'Display frequency tables'.

- * Three basic measures of central tendency, upper right part of screen: mean, median, and mode.
- * Three basic measures of variability, bottom left part of screen: variance, Standard Deviation, and range.
- * Skewness [Note. Skewness refers to the extent to which the data are normally distributed around the mean. Skewed data involve having either mostly high scores with a few low ones or having mostly low scores with a few high ones.] Readers are referred to the following sources for a more detailed definition of skewness: http://www.statistics.com/index.php?page=glossary&term_id=356¹⁴ and <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>¹⁵
- * Kurtosis [Note. Kurtosis also refers to the extent to which the data are normally distributed around the mean. This time, the data are piled up higher than normal around the mean or piled up higher than normal at the ends of the distribution.] Readers are referred to the following sources for a more detailed definition of kurtosis: http://www.statistics.com/index.php?page=glossary&term_id=326¹⁶ and <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>¹⁷

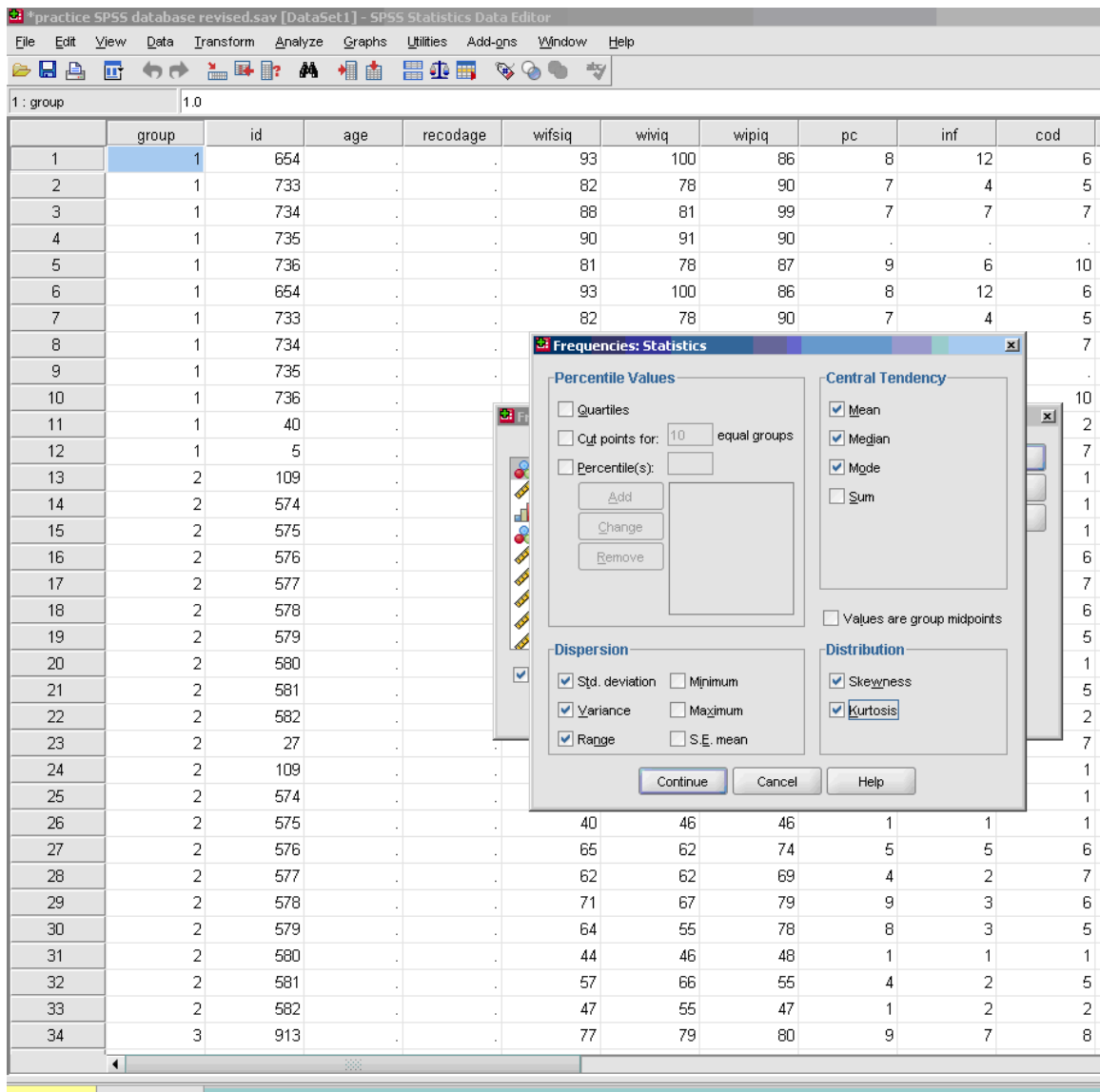
¹³<http://cnx.org/content/m37276/latest/figure1.3.PNG/image>

¹⁴http://www.statistics.com/index.php?page=glossary&term_id=356

¹⁵<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

¹⁶http://www.statistics.com/index.php?page=glossary&term_id=326

¹⁷<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>



18

* Charts (optional, use only if you want a visual depiction of your data)

* Histograms (optional, use only if you want a visual depiction of your data) with normal curve

¹⁸<http://cnx.org/content/m37276/latest/figure1.4.PNG/image>

TAKS_2009.sav [DataSet1] - SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

23:

	CAMPUS	CH004TR09 R	CL004TR09 R	CW004TR09 R	var	var	var	var	var	var	var
1	1902001	.	.	.							
2	1902041	.	.	.							
3	1902103	.	.	97							
4	1903001	.	.	.							
5	1903041	.	.	.							
6	1903101	89	.	97							
7	1904001	.	.	.							
8	1904041	.	.	.							
9	1904102	.	.	87							
10	1906002	.	.	.							
11	1906102	.	.	88							
12	1907001	.	.	.							
13	1907041	.	.	.							
14	1907101	.	.	.							
15	1907107	82	33	91							
16	1907110	82	33	91							
17	1908002	.	.	.							
18	1908041	.	.	.							
19	1908101	57	.	90							
20	1908102	57	.	90							
21	1909001	.	.	.							
22	1909101	.	.	70							
23	2901001	.	.	.							
24	2901002	.	.	.							
25	2901041	.	.	.							
26	2901101	.	.	.							
27	2901104	79	38	89							
28	2901105	.	.	.							

Frequencies

CAMPUS

Display frequency tables

Frequencies: Charts

Chart Type

☒ None

☐ Bar charts

☐ Pie charts

☐ Histograms:

☐ With normal curve

Chart Values

☒ Frequencies ☐ Percentages

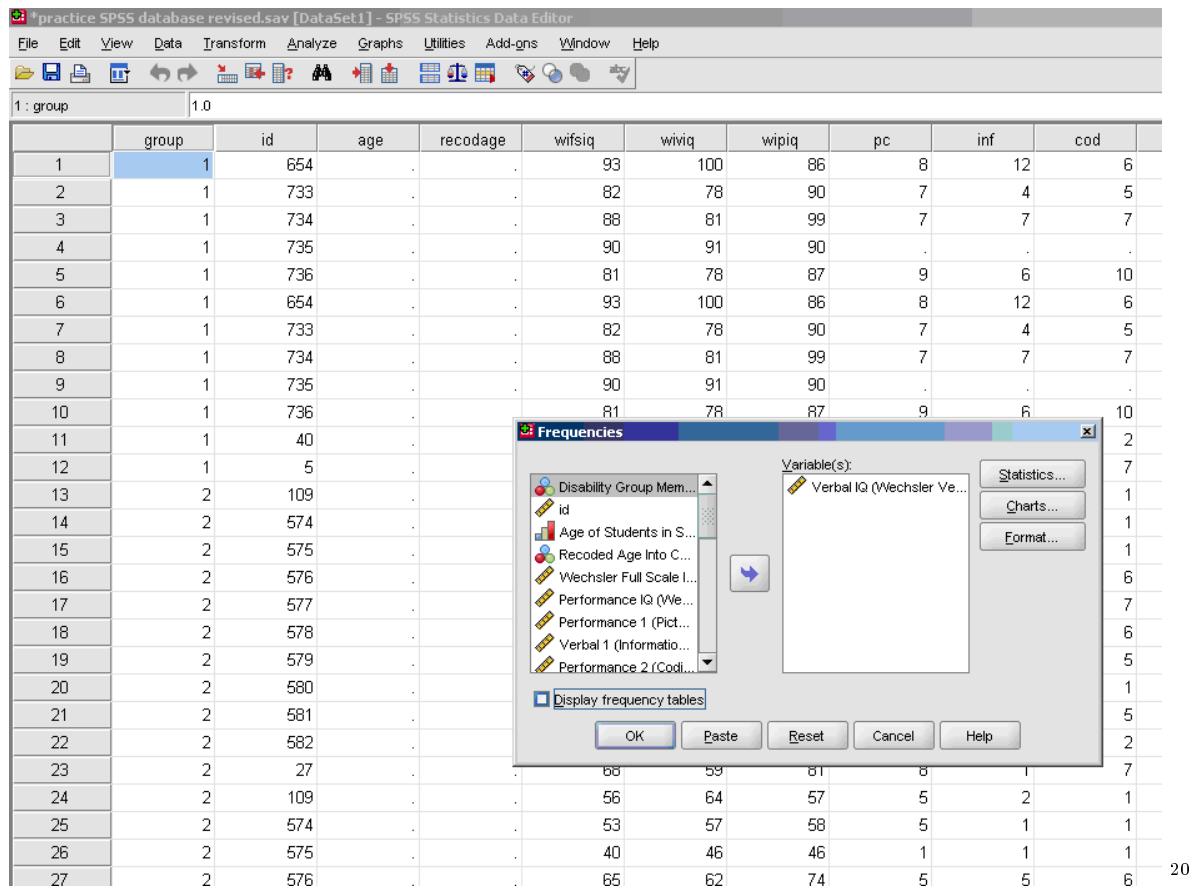
Continue Cancel Help

Statistics... Charts... Format...

19

- * Uncheck the display frequency tables so that you are not provided with the frequencies of your data every time descriptive statistics are obtained.
- * OK

¹⁹<http://cnx.org/content/m37276/latest/figure1.5.PNG/image>



The screenshot shows the SPSS Statistics Data Editor window with a dataset named 'practice SPSS database revised.sav [DataSet1]'. The dataset has 27 rows and 10 columns: group, id, age, recodage, wfsiq, wviq, wpiq, pc, inf, and cod. The 'group' column has two values: 1 and 2. The 'id' column contains unique identifiers for each row. The 'age' column contains ages ranging from 27 to 78. The 'recodage' column contains values 1 and 2. The 'wfsiq' column contains values ranging from 40 to 93. The 'wviq' column contains values ranging from 59 to 100. The 'wpiq' column contains values ranging from 57 to 99. The 'pc' column contains values ranging from 5 to 9. The 'inf' column contains values ranging from 4 to 12. The 'cod' column contains values ranging from 6 to 10. A 'Frequencies' dialog box is open, showing the 'Verbal IQ (Wechsler Ve...' variable selected in the 'Variable(s)' list. The dialog also has buttons for 'Statistics...', 'Charts...', and 'Format...'. The 'Display frequency tables' checkbox is checked.

To obtain descriptive statistics for subgroups, do the following:

* Split File (icon middle top of screen next to the scales)

²⁰<http://cnx.org/content/m37276/latest/figure1.6.PNG/image>

***practice SPSS database revised.sav [DataSet1] - SPSS Statistics Data Editor**

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

1: group 1.0 Split File

	group	id	age	recodage	wifsiq	wiviq	wipiq	pc
1	1	654	.	.	93	100	86	
2	1	733	.	.	82	78	90	
3	1	734	.	.	88	81	99	
4	1	735	.	.	90	91	90	
5	1	736	.	.	81	78	87	
6	1	654	.	.	93	100	86	
7	1	733	.	.	82	78	90	
8	1	734	.	.	88	81	99	
9	1	735	.	.	90	91	90	
10	1	736	.	.	81	78	87	
11	1	40	.	.	86	88	87	1
12	1	5	.	.	83	85	83	1
13	2	109	.	.	56	64	57	
14	2	574	.	.	53	57	58	
15	2	575	.	.	40	46	46	
16	2	576	.	.	65	62	74	
17	2	577	.	.	62	62	69	
18	2	578	.	.	71	67	79	
19	2	579	.	.	64	55	78	
20	2	580	.	.	44	46	48	
21	2	581	.	.	57	66	55	
22	2	582	.	.	47	55	47	
23	2	27	.	.	68	59	81	
24	2	109	.	.	56	64	57	
25	2	574	.	.	53	57	58	
26	2	575	.	.	40	46	46	
27	2	576	.	.	65	62	74	
28	2	577	.	.	62	62	69	
29	2	578	.	.	71	67	79	
30	2	579	.	.	64	55	78	
31	2	580	.	.	44	46	48	
32	2	581	.	.	57	66	55	
33	2	582	.	.	47	55	47	
34	3	913	.	.	77	79	80	

Data View Variable View

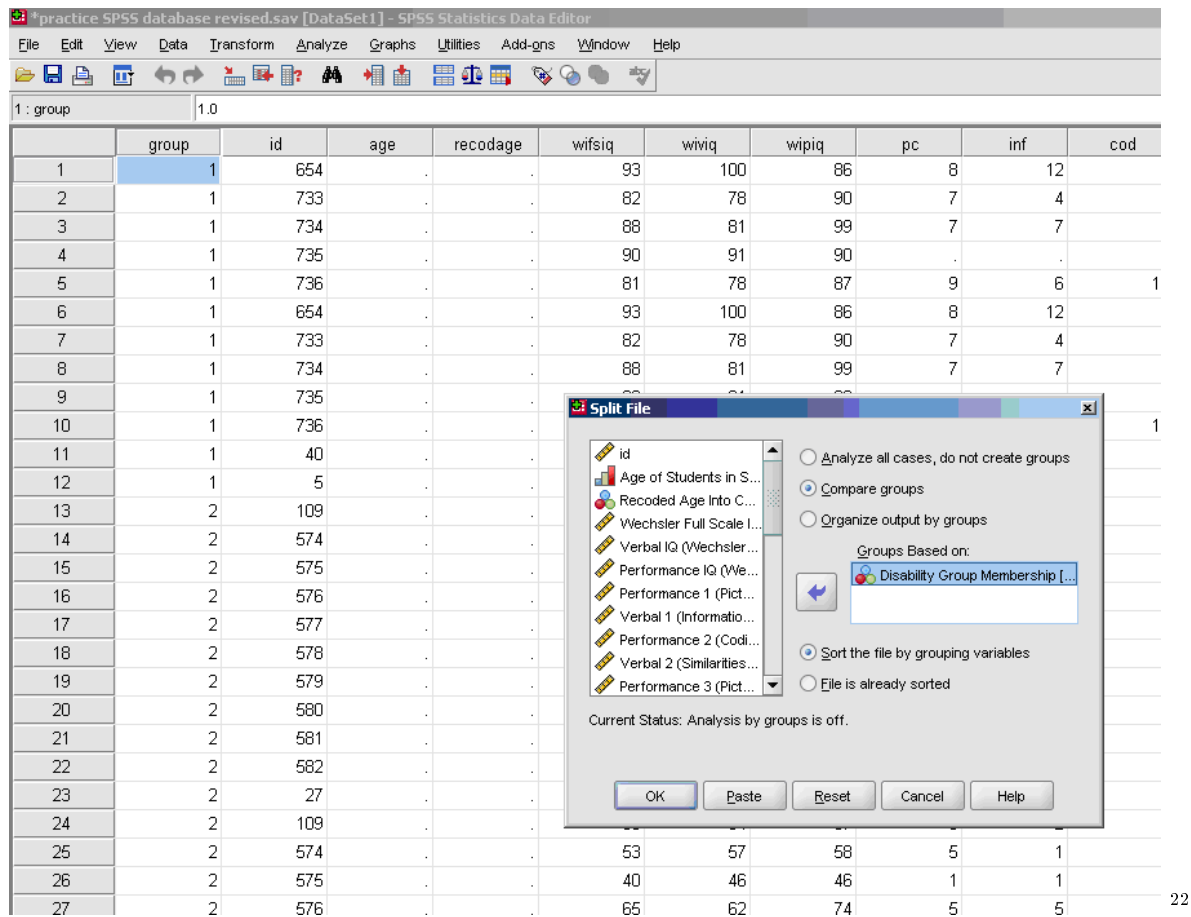
Split File

Start Inbox - Microsoft Outlook *practice SPSS databa... *Output2 [Document2] - ... stepstorunningdescriptive...

* Compare Groups

* Click on group (typically dichotomous in nature) and move to empty cell.

²¹<http://cnx.org/content/m37276/latest/figure1.7.PNG/image>



The screenshot shows the SPSS Statistics Data Editor window with a dataset named 'practice SPSS database revised.sav [DataSet1]'. The dataset has 27 rows and 10 columns: group, id, age, recodage, wfsiq, wiviq, wipiq, pc, inf, and cod. A 'Split File' dialog box is open, showing the following options:

- ☐ Analyze all cases, do not create groups
- ☒ Compare groups
- ☐ Organize output by groups

Under 'Groups Based on:', the variable 'Disability Group Membership [...]' is selected. The 'Sort the file by grouping variables' option is also selected. The 'Current Status' is 'Analysis by groups is off.' The dialog box has buttons for OK, Paste, Reset, Cancel, and Help.

* OK

✓ Analyze

* Descriptive Statistics

* Frequencies

* Move over the dependent variable/s

* Do **NOT** move over the independent variable/s or any string variables

²²<http://cnx.org/content/m37276/latest/figure1.8.PNG/image>

*practice SPSS database revised.sav [DataSet1] - SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

1 : group 1.0

	group	id	age	recodage	wfsiq	wiviq	wipiq	pc	inf	cod
1	1	654	.	.	93	100	86	8	12	6
2	1	733	.	.	82	78	90	7	4	5
3	1	734	.	.	88	81	99	7	7	7
4	1	735	.	.	90	91	90	.	.	.
5	1	736	.	.	81	78	87	9	6	10
6	1	654	.	.	93	100	86	8	12	6
7	1	733	.	.	82	78	90	7	4	5
8	1	734	.	.	88	81	99	7	7	7
9	1	735	.	.	90	91	90	.	.	.
10	1	736	.	.	81	78	87	9	6	10
11	1	40	.	.	.	.	.	.	.	.
12	1	5	.	.	.	.	.	.	.	.
13	1	614	6	1.00	.	.	.	.	.	.
14	1	624	6	1.00	.	.	.	.	.	.
15	1	628	6	1.00	.	.	.	.	.	.
16	1	635	6	1.00	.	.	.	.	.	.
17	1	664	6	1.00	.	.	.	.	.	.
18	1	614	6	1.00	.	.	.	.	.	.
19	1	624	6	1.00	.	.	.	.	.	.
20	1	628	6	1.00	.	.	.	.	.	.
21	1	635	6	1.00	.	.	.	.	.	.
22	1	664	6	1.00	.	.	.	.	.	.
23	1	522	6	1.00	71	70	78	7	6	1
24	1	524	6	1.00	84	92	79	4	5	10
25	1	616	6	1.00	76	75	80	6	7	17

Frequencies

Variable(s): Verbal IQ (Wechsler Ve... Statistics... Charts... Format...

Display frequency tables

OK Paste Reset Cancel Help

✓ Statistics

- * Three basic measures of central tendency, upper right part of screen: mean, median, and mode.
- * Three basic measures of variability, bottom left part of screen: variance, Standard Deviation, and range.
- * Skewness [Note. Skewness refers to the extent to which the data are normally distributed around the mean. Skewed data involve having either mostly high scores with a few low ones or having mostly low scores with a few high ones.] Readers are referred to the following sources for a more detailed definition of skewness: http://www.statistics.com/index.php?page=glossary&term_id=356²⁴ and <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>²⁵
- * Kurtosis [Note. Kurtosis also refers to the extent to which the data are normally distributed around the mean. This time, the data are piled up higher than normal around the mean or piled up higher than normal at the ends of the distribution.] Readers are referred to the following sources for a more detailed definition of kurtosis: http://www.statistics.com/index.php?page=glossary&term_id=326²⁶ and <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>²⁷

* Continue

²³<http://cnx.org/content/m37276/latest/figure1.9.PNG/image>

²⁴http://www.statistics.com/index.php?page=glossary&term_id=356

²⁵<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

²⁶http://www.statistics.com/index.php?page=glossary&term_id=326

²⁷<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

1 : group 1.0

	group	id	age	recodage	wfsiq	wviq	wipiq	pc	inf	cod
1	1	654	.	.	93	100	86	8	12	6
2	1	733	.	.	82	78	90	7	4	5
3	1	734	.	.	88	81	99	7	7	7
4	1	735	.	.	90	91	90	.	.	.
5	1	736	.	.	81	78	87	9	6	10
6	1	654	.	.	93	100	86	8	12	6
7	1	733	.	.	82	78	90	7	4	5
8	1	734	.	.	.	.	.	.	.	.
9	1	735	.	.	.	.	.	.	.	.
10	1	736	.	.	.	.	.	.	.	.
11	1	40	.	.	.	.	.	.	.	.
12	1	5	.	.	.	.	.	.	.	.
13	1	614	6	1.00	.	.	.	.	.	.
14	1	624	6	1.00	.	.	.	.	.	.
15	1	628	6	1.00	.	.	.	.	.	.
16	1	635	6	1.00	.	.	.	.	.	.
17	1	664	6	1.00	.	.	.	.	.	.
18	1	614	6	1.00	.	.	.	.	.	.
19	1	624	6	1.00	.	.	.	.	.	.
20	1	628	6	1.00	.	.	.	.	.	.
21	1	635	6	1.00	.	.	.	.	.	.
22	1	664	6	1.00	.	.	.	.	.	.
23	1	522	6	1.00	.	.	.	.	.	.
24	1	524	6	1.00	.	.	.	.	.	.
25	1	619	6	1.00	.	.	.	.	.	.
26	1	123	7	1.00	80	76	87	5	7	17
27	1	278	7	1.00	105	101	108	10	9	10

28

* Charts (optional, use only if you want a visual depiction of your data)

* Histograms (optional, use only if you want a visual depiction of your data)with normal curve

²⁸<http://cnx.org/content/m37276/latest/figure1.10.PNG/image>

The screenshot shows the SPSS Statistics Data Editor interface. The main window displays a dataset with the following columns: group, id, age, recodage, wifsiq, wiviq, wipiq, pc, inf, cod. The data is organized into rows, with the first row highlighted. A dialog box titled 'Frequencies: Charts' is open, showing the 'Chart Type' section with 'Histograms' selected and 'With normal curve' checked. The 'Chart Values' section has 'Frequencies' selected. The 'Continue' button is highlighted.

* OK

To calculate a z -score for any continuous variable:

✓ Analyze

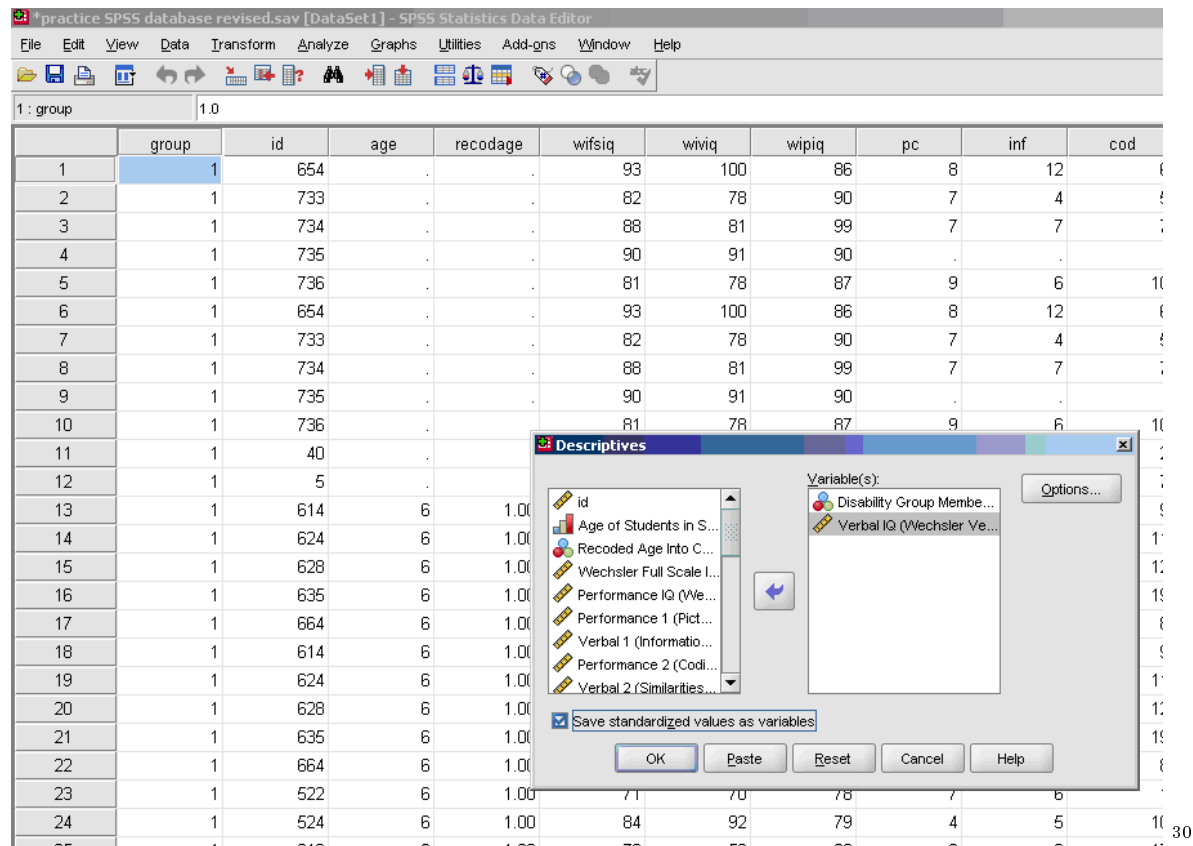
* Descriptive Statistics

* Descriptives

* Send variable on which you want z -scores to be calculated to empty cell

* Check box for Save standardized values as variables

²⁹<http://cnx.org/content/m37276/latest/figure1.11.PNG/image>



* OK

* You will be sent to the output window, as shown in Table 1. [Note. In some versions of SPSS, you will not be sent to the output window, but will remain in the data window.] The information in the output window is not relevant for your purposes. To see the variable that was just created, go to the SPSS data. The far right column should now be the new z-score variable that was created.

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
Verbal IQ (Wechsler Verbal Intelligence 3)	1182	46	129	77.97	13.661
Valid N (listwise)	1182				

Table 2.1: Descriptive Statistics

* A new variable/s will have been generated for you in the data window

To get this information in a usable output form, do the following:

✓ Analyze

* Descriptive Statistics

* Frequencies

* Move over the newly created z-score variable(s) (z-scores will generally appear at the bottom of your list with the words: "Zscore: Verbal IQ (Wechsler Verbal Intelligence 3)

³⁰<http://cnx.org/content/m37276/latest/figure1.12.PNG/image>

- * Make sure the frequencies box is checked
- * OK

- * Copy or cut the frequency table for this z-score variable and carry it into WORD. Delete any irrelevant information.

		Zscore: Verbal IQ (Wechsler Verbal Intelligence 3)	Zscore(wiviq) Verbal IQ (Wechsler Verbal Intelligence 3)
N	Valid	1182	1182
	Missing	0	0

Table 2.2: Z Scores

To calculate a *T*-score for any continuous variable:

- ✓ Analyze
- * Descriptive Statistics
- * Descriptives
- * Send variable on which you want *T* scores to be calculated to empty cell
- * Check box for Save standardized values as variables

The screenshot displays the SPSS Statistics Data Editor window. The main data grid shows 25 cases with columns: group, id, age, recodage, wfsiq, wiviq, wpiq, pc, inf, cod. A 'Descriptives' dialog box is open in the foreground, with 'Verbal IQ (Wechsler Ve...' selected as the variable for analysis. The 'Save standardized values as variables' checkbox is checked. The background data is as follows:

Case	group	id	age	recodage	wfsiq	wiviq	wpiq	pc	inf	cod
1	1	654	.	.	93	100	86	8	12	6
2	1	733	.	.	82	78	90	7	4	5
3	1	734	.	.	88	81	99	7	7	7
4	1	735	.	.	90	91	90	.	.	.
5	1	736	.	.	81	78	87	9	6	10
6	1	654	.	.	93	100	86	8	12	6
7	1	733	.	.	82	78	90	7	4	5
8	1	734	.	.	88	81	99	7	7	7
9	1	735	.	.	90	91	90	.	.	.
10	1	736	.	.	81	78	87	9	6	10
11	1	40	.	.	.	.	.	.	.	.
12	1	5	.	.	.	.	.	.	.	.
13	1	614	6	1.00	.	.	.	.	.	.
14	1	624	6	1.00	.	.	.	.	.	.
15	1	628	6	1.00	.	.	.	.	.	.
16	1	635	6	1.00	.	.	.	.	.	.
17	1	664	6	1.00	.	.	.	.	.	.
18	1	614	6	1.00	.	.	.	.	.	.
19	1	624	6	1.00	.	.	.	.	.	.
20	1	628	6	1.00	.	.	.	.	.	.
21	1	635	6	1.00	.	.	.	.	.	.
22	1	664	6	1.00	.	.	.	.	.	.
23	1	522	6	1.00	71	70	76	7	6	1
24	1	524	6	1.00	84	92	79	4	5	10
25	4	616	6	1.00	76	78	88	6	5	17

³¹ <http://cnx.org/content/m37276/latest/figure1.13.PNG/image>

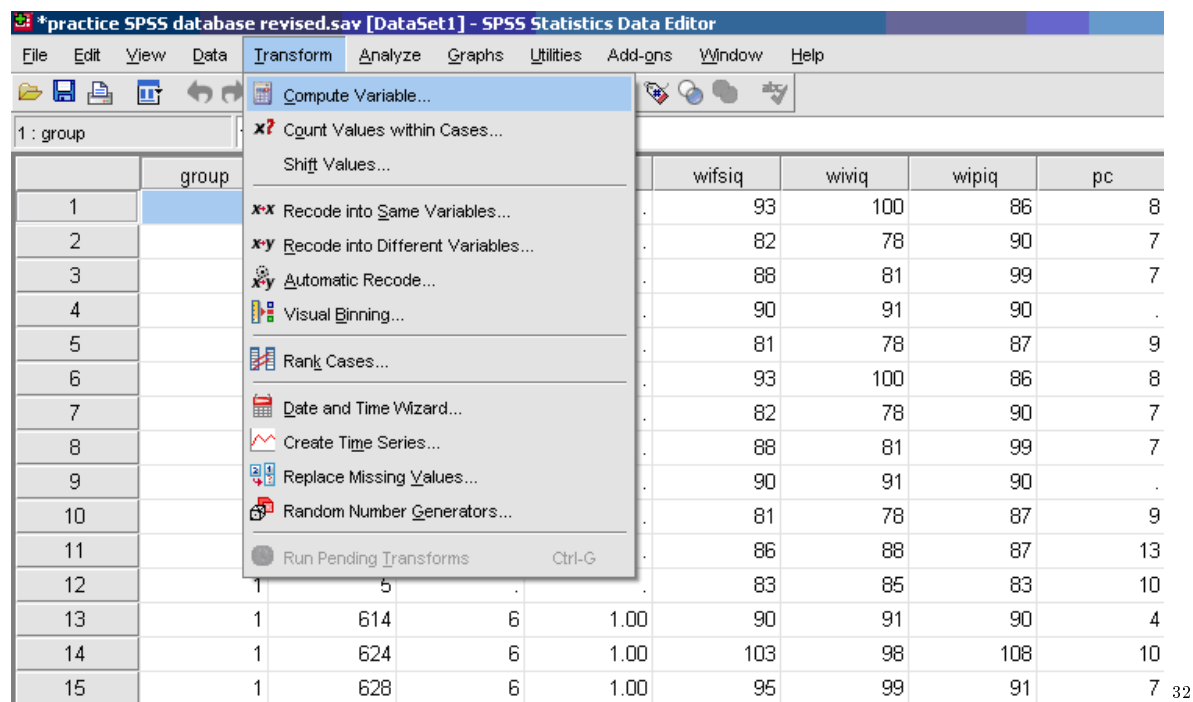
- * OK
- * You will be sent to the output window. Nothing in the output window is helpful. Go to the SPSS data screen by clicking on the data button bottom of screen. A new variable(s) will have been generated for you. This variable will be inserted into a formula so that you can have T scores.
- * Variable view window

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
Verbal IQ (Wechsler Verbal Intelligence 3)	1182	46	129	77.97	13.661
Valid N (listwise)	1182				

Table 2.3: Descriptive Statistics

✓ Create a new variable for your T score variable

- * Data view window
- * Transform
- * Compute Variable



- * Name your target variable the name you just generated for your T score variable
- * In the numeric expression window, type:
- * $50 + (10 \times [\text{name of the z-score variable generated by the computer earlier}])$

³²<http://cnx.org/content/m37276/latest/figure1.14.PNG/image>

SPSS Statistics Data Editor window showing a dataset with 30 rows and 9 columns. The columns are: TVerbalIQ, Zp, Zwiq, ZSco01, ZSco02, StdZ01, StdZ02, ZzZz01, and ZqZq01. The data is as follows:

	TVerbalIQ	Zp	Zwiq	ZSco01	ZSco02	StdZ01	StdZ02	ZzZz01	ZqZq01
1	1.54412	.	-1.05208	1.61281	-1.05208	1.61281	1.61281	1.61281	1.61281
2	-0.60798	.	-1.05208	0.00242	-1.05208	0.00242	0.00242	0.00242	0.00242
3	-0.31451	.	-1.05208	0.22201	-1.05208	0.22201	0.22201	0.22201	0.22201
4	0.66372	.	-1.05208	0.95401	-1.05208	0.95401	0.95401	0.95401	0.95401
5	-0.60798	.	-1.05208	0.00242	-1.05208	0.00242	0.00242	0.00242	0.00242
6	1.54412	.	-1.05208	1.61281	-1.05208	1.61281	1.61281	1.61281	1.61281
7	-0.60798	.	-1.05208	0.00242	-1.05208	0.00242	0.00242	0.00242	0.00242
8	-0.31451	.	-1.05208	0.22201	-1.05208	0.22201	0.22201	0.22201	0.22201
9	0.66372	.	-1.05208	0.95401	-1.05208	0.95401	0.95401	0.95401	0.95401
10	-0.60798	.	-1.05208	0.00242	-1.05208	0.00242	0.00242	0.00242	0.00242
11	0.37025	.	-1.05208	0.22201	-1.05208	0.22201	0.22201	0.22201	0.22201
12	0.07678	.	-1.05208	0.95401	-1.05208	0.95401	0.95401	0.95401	0.95401
13	0.66372	.	-1.05208	0.00242	-1.05208	0.00242	0.00242	0.00242	0.00242
14	1.34848	.	-1.05208	1.61281	-1.05208	1.61281	1.61281	1.61281	1.61281
15	1.44630	.	-1.05208	0.00242	-1.05208	0.00242	0.00242	0.00242	0.00242
16	1.34848	.	-1.05208	0.22201	-1.05208	0.22201	0.22201	0.22201	0.22201
17	0.85936	.	-1.05208	0.95401	-1.05208	0.95401	0.95401	0.95401	0.95401
18	0.66372	.	-1.05208	0.00242	-1.05208	0.00242	0.00242	0.00242	0.00242
19	1.34848	.	-1.05208	1.61281	-1.05208	1.61281	1.61281	1.61281	1.61281
20	1.44630	.	-1.05208	0.00242	-1.05208	0.00242	0.00242	0.00242	0.00242
21	1.34848	.	-1.05208	0.22201	-1.05208	0.22201	0.22201	0.22201	0.22201
22	0.85936	.	-1.05208	0.95401	-1.05208	0.95401	0.95401	0.95401	0.95401
23	-1.39057	.	-1.05208	0.00242	-1.05208	0.00242	0.00242	0.00242	0.00242
24	0.76154	.	-1.05208	1.61281	-1.05208	1.61281	1.61281	1.61281	1.61281
25	-2.56444	.	-1.05208	0.00242	-1.05208	0.00242	0.00242	0.00242	0.00242
26	-0.80363	.	-1.05208	0.22201	-1.05208	0.22201	0.22201	0.22201	0.22201
27	1.64194	.	-1.05208	0.95401	-1.05208	0.95401	0.95401	0.95401	0.95401
28	-0.11887	.	-1.05208	0.00242	-1.05208	0.00242	0.00242	0.00242	0.00242
29	-0.80363	.	-1.05208	1.61281	-1.05208	1.61281	1.61281	1.61281	1.61281
30	-0.51016	.	-1.05208	0.00242	-1.05208	0.00242	0.00242	0.00242	0.00242

The 'Compute Variable' dialog box is open, showing the target variable 'TVerbalIQ' and the numeric expression $50 + 10 * (\text{StdZ02})$. The function group is set to 'Arithmetic'.

* OK

³³<http://cnx.org/content/m37276/latest/figure1.15.PNG/image>

*practice SPSS database revised.sav [DataSet1] - SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

1 : TVerbalQ 66.12809669195829

	oup	Zwiviq	ZSco01	ZSco02	StdZ01	StdZ02	ZzZz01
1	.	1.54412	-1.05208	1.61281	-1.05208	1.61281	1.61281
2	.	-0.60798	-1.05208	0.00242	-1.05208	0.00242	0.00242
3	.	-0.31451	-1.05208	0.22201	-1.05208	0.22201	0.22201
4	.	0.66372	-1.05208	0.95401	-1.05208	0.95401	0.95401
5	.	-0.60798	-1.05208	0.00242	-1.05208	0.00242	0.00242
6	.	1.54412	-1.05208	1.61281	-1.05208	1.61281	1.61281
7	.	-0.60798	-1.05208	0.00242	-1.05208	0.00242	0.00242
8	.	-0.31451	-1.05208	0.22201	-1.05208	0.22201	0.22201
9	.	0.66372	-1.05208	0.95401	-1.05208	0.95401	0.95401
10	.	-0.60798	-1.05208	0.00242	-1.05208	0.00242	0.00242
11	.	0.37025	-1.05208	0.73441	-1.05208	0.73441	0.73441
12	.	0.07678	-1.05208	0.51481	-1.05208	0.51481	0.51481
13	.	0.66372	-1.05208	0.95401	-1.05208	0.95401	0.95401
14	.	1.34848	-1.05208	1.46641	-1.05208	1.46641	1.46641
15	.	1.44630	-1.05208	1.53961	-1.05208	1.53961	1.53961
16	.	1.34848	-1.05208	1.46641	-1.05208	1.46641	1.46641
17	.	0.85936	-1.05208	1.10041	-1.05208	1.10041	1.10041
18	.	0.66372	-1.05208	0.95401	-1.05208	0.95401	0.95401
19	.	1.34848	-1.05208	1.46641	-1.05208	1.46641	1.46641
20	.	1.44630	-1.05208	1.53961	-1.05208	1.53961	1.53961
21	.	1.34848	-1.05208	1.46641	-1.05208	1.46641	1.46641
22	.	0.85936	-1.05208	1.10041	-1.05208	1.10041	1.10041
23	.	-1.39057	-1.05208	-0.58318	-1.05208	-0.58318	-0.58318
24	.	0.76154	-1.05208	1.02721	-1.05208	1.02721	1.02721
25	.	-2.56444	-1.05208	-1.46158	-1.05208	-1.46158	-1.46158
26	.	-0.80363	-1.05208	-0.14398	-1.05208	-0.14398	-0.14398
27	.	1.64194	-1.05208	1.68601	-1.05208	1.68601	1.68601
28	.	-0.11887	-1.05208	0.36841	-1.05208	0.36841	0.36841
29	.	-0.80363	-1.05208	-0.14398	-1.05208	-0.14398	-0.14398
30	.	-0.51016	-1.05208	0.07561	-1.05208	0.07561	0.07561
31	.	-0.21669	-1.05208	0.29521	-1.05208	0.29521	0.29521
32	.	0.27242	-1.05208	0.66121	-1.05208	0.66121	0.66121
33	.	-0.41234	-1.05208	0.14881	-1.05208	0.14881	0.14881
34	.	-0.11887	-1.05208	0.36841	-1.05208	0.36841	0.36841

34

- * Respond yes to change existing variable
- * You may be sent to the output screen. Nothing there is helpful.
- * Go to data button and view your new T score variable.
- * To get this information in a usable output form, do the following:

✓ Analyze

- * Descriptive Statistics
- * Frequencies
- * Move over the newly created T score variable
- * Make sure the frequencies box is checked.

³⁴<http://cnx.org/content/m37276/latest/figure1.16.PNG/image>

The screenshot shows the SPSS Statistics Data Editor window with a dataset named 'practice SPSS database revised.sav [DataSet1]'. The dataset has 32 rows and 7 columns: Group, Zwiwig, ZSco01, ZSco02, StdZ01, StdZ02, and ZzZ01. A 'Frequencies' dialog box is open, showing 'TVerbalQ' as the variable selected for frequency analysis. The dialog box includes options for 'Statistics...', 'Charts...', and 'Format...', and a checkbox for 'Display frequency tables' which is checked.

* OK

* Copy or cut the frequency table for this *T* score variable and carry it into WORD. Delete any irrelevant information.

2.3 Writing Up Your Statistics

So, how do you "write up" your Research Questions and your Results? Schuler W. Huck (2000) in his seminal book entitled, *Reading Statistics and Research*, points to the importance of your audience understanding and making sense of your research in written form. Huck further states:

2.3.1

This book is designed to help people *decipher* what researchers are trying to communicate in the written or oral summaries of their investigations. Here, the goal is simply to distill meaning from the words, symbols, tables, and figures included in the research report. To be competent in this arena, one must not only be able to decipher what's presented but also to "fill in the holes"; this is the case because researchers typically

³⁵<http://cnx.org/content/m37276/latest/figure1.17.PNG/image>

assume that those receiving the research report are familiar with unmentioned details of the research process and statistical treatment of data.

A Note from the Editors

Researchers and Professors John Slate and Ana Rojas-LeBouef understand this critical issue, so often neglected or not addressed by other authors and researchers. They point to the importance of doctoral students "writing up their statistics" in a way that others can understand your reporting and as importantly, interpret the meaning of your significant findings and implications for the preparation and practice of educational leadership. Slate and LeBouef provide you with a model for "writing up your descriptive statistics."

Click here to view: *Writing Up Your Descriptive Statistics* ³⁶

2.4 References

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum

Hyperstats Online Statistics Textbook. (n.d.) Retrieved from <http://davidmlane.com/hyperstat/> ³⁷

Kurtosis. (n.d.). Definition. Retrieved from http://www.statistics.com/index.php?page=glossary&term_id=326 ³⁸

Kurtosis. (n.d.). *Definition of normality*. Retrieved from <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb> ³⁹

Onwuegbuzie, A. J., & Daniel, L. G. (2002). Uses and misuses of the correlation coefficient. *Research in the Schools*, 9(1), 73-90.

Skewness. (n.d.) Retrieved from http://www.statistics.com/index.php?page=glossary&term_id=356 ⁴⁰

Skewness. (n.d.). *Definition of normality*. Retrieved from <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb> ⁴¹

StatSoft, Inc. (2011). *Electronic statistics textbook*. Tulsa, OK: StatSoft. WEB: <http://www.statsoft.com/textbook/> ⁴²

³⁶<http://cnx.org/content/col11299/latest/>

³⁷<http://davidmlane.com/hyperstat/>

³⁸http://www.statistics.com/index.php?page=glossary&term_id=326

³⁹<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

⁴⁰http://www.statistics.com/index.php?page=glossary&term_id=356

⁴¹<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

⁴²<http://www.statsoft.com/textbook/>

Chapter 3

Calculating a Nonparametric Pearson Chi-Square



NOTE: This chapter has been peer-reviewed, accepted, and endorsed by the National Council of Professors of Educational Administration (NCPEA) as a significant contribution to the scholarship and practice of education administration. Formatted and edited in Connexions by Theodore Creighton and Brad Bizzell, Virginia Tech, Janet Tareilo, Stephen F. Austin State University, and Thomas Kersten, Roosevelt University.

3.1

This chapter is part of a larger Collection (Book) and is available at: *Calculating Basic Statistical Procedures in SPSS: A Self-Help and Practical Guide to Preparing Theses, Dissertations, and Manuscripts* ²

NOTE: Slate and LeBouef have written a "companion book" which is available at: *Preparing and Presenting Your Statistical Findings: Model Write Ups* ³

Authors Information

John R. Slate is a Professor at Sam Houston State University where he teaches Basic and Advanced Statistics courses, as well as professional writing, to doctoral students in Educational Leadership and Counseling. His research interests lie in the use of educational databases, both state and national, to reform school practices. To date, he has chaired and/or served over 100 doctoral student dissertation committees. Recently, Dr. Slate created a website (Writing and Statistical Help⁴) to assist students and faculty with both statistical assistance and in editing/writing their dissertations/theses and manuscripts.

¹This content is available online at <<http://cnx.org/content/m37277/1.5/>>.

²<http://my.qoop.com/store/NCPEA-Publications-1781472103076212/>

³<http://my.qoop.com/store/NCPEA-Publications-1781472103076212/>

⁴<http://cnx.org/content/m37277/latest/www.writingandstatisticalhelp>

Ana Rojas-LeBoeuf is a Literacy Specialist at the Reading Center at Sam Houston State University where she teaches developmental reading courses. Dr. LeBoeuf recently completed her doctoral degree in Reading, where she conducted a 16-year analysis of Texas statewide data regarding the achievement gap. Her research interests lie in examining the inequities in achievement among ethnic groups. Dr. Rojas-LeBoeuf also assists students and faculty in their writing and statistical needs on the Writing and Statistical Help website.

Editors Information

Theodore B. Creighton, is a Professor at Virginia Tech and the Publications Director for NCPEA Publications⁵, the Founding Editor of Education Leadership Review,⁶ and the Senior Editor of the NCPEA Connexions Project.

Brad E. Bizzell, is a recent graduate of the Virginia Tech Doctoral Program in Educational Leadership and Policy Studies, and is a School Improvement Coordinator for the Virginia Tech Training and Technical Assistance Center. In addition, Dr. Bizzell serves as an Assistant Editor of the NCPEA Connexions Project in charge of technical formatting and design.

Janet Tareilo, is a Professor at Stephen F. Austin State University and serves as the Assistant Director of NCPEA Publications. Dr. Tareilo also serves as an Assistant Editor of the NCPEA Connexions Project and as a editor and reviewer for several national and international journals in educational leadership.

Thomas Kersten is a Professor at Roosevelt University in Chicago. Dr. Kersten is widely published and an experienced editor and is the author of Taking the Mystery Out of Illinois School Finance⁷, a Connexions Print on Demand publication. He is also serving as Editor in Residence for this book by Slate and LeBoeuf.

3.2 Conducting a Nonparametric Pearson Chi-Square

In this set of steps, readers are provided with directions on calculating a statistical procedure in which the independent variable and the dependent variable are categorical variables. As such, the only descriptive statistics that can be obtained are frequencies, percentages, and sums. Because the data on which this chi-square procedure is used are grouped data, skewness and kurtosis values are not appropriate. Readers should ensure that the assumptions described in the steps below are met prior to conducting this nonparametric procedure. For more detailed information about the statistical and conceptual underpinnings of this statistical technique, readers are referred to the Hyperstats Online Statistics Textbook at http://davidmlane.com/hyperstat/chi_square.html⁸ or to the *Electronic Statistics Textbook* (2011) at <http://www.statsoft.com/textbook/basic-statistics/>⁹

3.2.1 Step One:

Check to make sure that both variables are categorical in nature. That is, the variables must have values that are in a restricted range (e.g., 1 or 2 for gender; 1 – 5 for Strongly Agree through Strongly Disagree; 1 – 5 for ethnicity categories).

3.2.2 Step Two:

Check to verify that you have available per cell at least 5 responses (i.e., divide the sample size by the number of cells [number of categories for the IV times the number of categories for the DV] and have a value of at

⁵<http://www.ncpeapublications.org>

⁶<http://ncpeapublications.org/about-elr.html>

⁷<http://cnx.org/content/col10606/latest/>

⁸http://davidmlane.com/hyperstat/chi_square.html

⁹<http://www.statsoft.com/textbook/basic-statistics/>

least 5).

3.2.3 Step Three:

Verify that only one response per participant is present. Once these assumptions have been checked and validated, then the Pearson chi-square procedure can be calculated.

3.2.4 Step Four:

- ✓ Analyze
- * Descriptive Statistics
- * Crosstabs

The screenshot displays the SPSS Statistics Data Editor window for a file named '*practice SPSS database revised.sav [DataSet1]'. The 'Analyze' menu is open, showing various statistical options. The 'Crosstabs...' option is highlighted. The data table below shows two columns: 'pup' and 'Zwiv'. The 'pup' column contains values from 1 to 34, and the 'Zwiv' column contains values from -1.0520 to -1.0520. The bottom status bar indicates 'Crosstabs...' and the taskbar shows the file name '*practice SPSS databa...' and the task 'PEARSON CHI'.

	pup	Zwiv	StdZ01
1	.		
2	.		
3	.		
4	.		
5	.		
6	.		
7	.		
8	.		
9	.		
10	.		
11	.		
12	.		
13	.		
14	.		
15	.		
16	.		
17	.		
18	.		
19	.	1.34848	-1.05208
20	.	1.44630	-1.05208
21	.	1.34848	-1.05208
22	.	0.85936	-1.05208
23	.	-1.39057	-1.05208
24	.	0.76154	-1.05208
25	.	-2.56444	-1.05208
26	.	-0.80363	-1.05208
27	.	1.64194	-1.05208
28	.	-0.11887	-1.05208
29	.	-0.80363	-1.05208
30	.	-0.51016	-1.05208
31	.	-0.21669	-1.05208
32	.	0.27242	-1.05208
33	.	-0.41234	-1.05208
34	.	-0.11887	-1.05208

¹⁰<http://cnx.org/content/m37277/latest/figure2.1.PNG/image>

- ✓ Independent Variable (e.g., gender) in Row
- ✓ Dependent Variable (e.g., responses to a survey item) in Column

7 : StdZ02 0.002415220101313482

	group	id	age	recodage	wfsiq	wiviq	wipiq	pc	inf	cod
1	1	654	.	.	93	100	86	8	12	6
2	1	733	.	.	82	78	90	7	4	5
3	1	734	.	.	88	81	99	7	7	7
4	1	735	.	.	90	91	90	.	.	.
5	1	736	.	.	81	78	87	9	6	10
6	1	654	.	.	93	100	86	8	12	6
7	1	733	.	.	82	78	90	7	4	5
8	1	734	.	.						
9	1	735	.	.						
10	1	736	.	.						
11	1	40	.	.						
12	1	5	.	.						
13	1	614	6	1						
14	1	624	6	1						
15	1	628	6	1						
16	1	635	6	1						
17	1	664	6	1						
18	1	614	6	1						
19	1	624	6	1						
20	1	628	6	1						
21	1	635	6	1						
22	1	664	6	1						
23	1	522	6	1						
24	1	524	6	1						
25	1	619	6	1						
26	1	123	7	1.00	80	76	87	5	7	17
27	1	278	7	1.00	105	101	108	10	9	10
28	1	599	7	1.00	76	83	72	5	5	8
29	1	602	7	1.00	77	76	81	9	2	6

11

- ✓ Cells
- ✓ In the Percentages Box
- ✓ Row

¹¹<http://cnx.org/content/m37277/latest/figure2.2.PNG/image>

The screenshot shows the SPSS Crosstabs: Cell Display dialog box. The 'Counts' section has ☒ Observed and ☐ Expected. The 'Percentages' section has ☒ Row, ☐ Column, and ☐ Total. The 'Residuals' section has ☐ Unstandardized, ☐ Standardized, and ☐ Adjusted standardized. The 'Noninteger Weights' section has ☒ Round cell counts, ☐ Round case weights, ☐ Truncate cell counts, ☐ Truncate case weights, and ☐ No adjustments. The 'Continue' button is highlighted.

	group	id	age	recodage	wifsiq	wiviq	wipiq	pc	inf	cod
1	1	654	.	.	93	100	86	8	12	6
2	1	733	.	.	82	78	90	7	4	5
3	1	734	.	.	88	81	99	7	7	7
4	1	735	.	.	90	91	90	.	.	.
5	1	736	.	.	81	78	87	9	6	10
6	1	654	.	.	93	100	86	8	12	6
7	1	733	.	.	82	78	90	7	4	5
8	1	734	.	.						
9	1	735	.	.						
10	1	736	.	.						
11	1	40	.	.						
12	1	5	.	.						
13	1	614	6	1						
14	1	624	6	1						
15	1	628	6	1						
16	1	635	6	1						
17	1	664	6	1						
18	1	614	6	1						
19	1	624	6	1						
20	1	628	6	1						
21	1	635	6	1						
22	1	664	6	1						
23	1	522	6	1						
24	1	524	6	1						
25	1	619	6	1						
26	1	123	7	1.00	80	76	87	5	7	17
27	1	278	7	1.00	105	101	108	10	9	10
28	1	599	7	1.00	76	83	72	5	5	8
29	1	602	7	1.00	77	76	81	9	2	6
30	1	618	7	1.00	83	79	91	8	5	9
31	1	626	7	1.00	90	82	102	9	10	8
32	1	629	7	1.00	98	87	112	13	8	10
33	1	647	7	1.00	91	80	106	10	4	13
34	1	653	7	1.00	83	83	86	3	7	14

Taskbar: Start | Calendar - Microsoft Outl... | *practice SPSS databa... | StepstoRunningaChiSqua... | PEARSON CHI SQUARE_a... | (0 unread) Yahoo! Mail, a...

- ✓ Continue
- ✓ Statistics
- ✓ Chi Square
- ✓ Phi and Cramer's V

¹²<http://cnx.org/content/m37277/latest/figure2.3.PNG/image>

The screenshot shows the SPSS Crosstabs: Statistics dialog box. The 'Chi-square' checkbox is checked. Under the 'Nominal' section, 'Phi and Cramer's V' is checked. Under the 'Ordinal' section, 'Gamma' is checked. Under the 'Nominal by Interval' section, 'Eta' is checked. The 'Test common odds ratio equals' field is set to 1. The 'Continue' button is highlighted.

	group	id	age	recodage	wfsiq	wiviq	wipiq	pc	inf	coc
1	1	654	.	.	93	100	86	8	12	
2	1	733	.	.	82	78	90	7	4	
3	1	734	.	.	88	81	99	7	7	
4	1	735	.	.	90	91	90	.	.	
5	1	736	.	.	81	78	87	9	6	
6	1	654	.	.	93	100	86	8	12	
7	1	733	.	.	82	78	90	7	4	
8	1	734	.	.						
9	1	735	.	.						
10	1	736	.	.						
11	1	40	.	.						
12	1	5	.	.						
13	1	614	6	1						
14	1	624	6	1						
15	1	628	6	1						
16	1	635	6	1						
17	1	664	6	1						
18	1	614	6	1						
19	1	624	6	1						
20	1	628	6	1						
21	1	635	6	1						
22	1	664	6	1						
23	1	522	6	1						
24	1	524	6	1						
25	1	619	6	1						
26	1	123	7	1.00	80	76	87	5	7	
27	1	278	7	1.00	105	101	108	10	9	
28	1	599	7	1.00	76	83	72	5	5	
29	1	602	7	1.00	77	76	81	9	2	
30	1	618	7	1.00	83	79	91	8	5	
31	1	626	7	1.00	90	82	102	9	10	
32	1	629	7	1.00	98	87	112	13	8	
33	1	647	7	1.00	91	80	106	10	4	
34	1	653	7	1.00	83	83	86	3	7	

✓ Continue

✓ OK

3.2.5 Step Five:

Check for Statistical Significance

1. Go to the Chi-Square Test Box
2. Find Pearson Chi-Square row and Asymp. Sig. (2-sided) column cell

¹³<http://cnx.org/content/m37277/latest/figure2.4.PNG/image>

Chi-Square Tests

	Value	df	Asymp.Sig. (2-sided)
Pearson Chi-Square	833.549 ^a	118	.000
Likelihood Ratio	907.609	118	.000
Linear-by-Linear	16.845	1	.000
Association			
N of Valid Cases	1182		

Table 3.1

a. 81 cells (45.0%) have expected count less than 5. The minimum expected count is .23.

3.3 Step Six:**Check Effect Size**

1. Go to the Symmetric Measures Box
2. Find the Nominal by Nominal Cramer's V row and Value column cell
3. The effect size is there and must be related to Cohen (1998)

Small effect size = .10 (range of .10 to .299)

Medium effect size = .30 (range of .30 to .499)

Large effect size = .50 (range of .50 to 1.00)

NOTE: Cramer's V cannot be greater than 1.00

Symmetric Measures

	Value	Approx Sig.
Nominal by Phi	.840	.000
Nominal Cramer's V	.94	.000
N of Valid Cases	1182	

Table 3.2**3.4 Step Seven:****Numerical Sentence**

1. $X^2(df)_{sp} = {}_{sp}\text{Pearson Chi-Square/Value Cell}_{,sp} p_{sp} < {}_{sp}.001$

$X^2(1) = 833.55, p < .001$

[Note. The sp refers to a space being present where the sp is located.]

3.5 Step Eight:

1. Go to the IV by DV table (i.e., the one above the Chi-Square Tests table)
2. Examine the percentages to determine where the statistically significant differences are

3.6 Step Nine:

Narrative and Interpretation Outline

1. Let the reader know what statistical procedure was conducted.
2. Explain how the assumptions for this statistical procedure were met.
3. Report the results from the test
4. Interpret the findings

3.7 Writing Up Your Statistics

So, how do you "write up" your Research Questions and your Results? Schuler W. Huck (2000) in his seminal book entitled, *Reading Statistics and Research*, points to the importance of your audience understanding and making sense of your research in written form. Huck further states:

3.7.1

This book is designed to help people *decipher* what researchers are trying to communicate in the written or oral summaries of their investigations. Here, the goal is simply to distill meaning from the words, symbols, tables, and figures included in the research report. To be competent in this arena, one must not only be able to decipher what's presented but also to "fill in the holes"; this is the case because researchers typically assume that those receiving the research report are familiar with unmentioned details of the research process and statistical treatment of data.

A Note from the Editors

Researchers and Professors John Slate and Ana Rojas-LeBouef understand this critical issue, so often neglected or not addressed by other authors and researchers. They point to the importance of doctoral students "writing up their statistics" in a way that others can understand your reporting and as importantly, interpret the meaning of your significant findings and implications for the preparation and practice of educational leadership. Slate and LeBouef provide you with a model for "writing up your Chi-square statistics."

Click here to view: *Writing Up Your Chi-square Statistics* ¹⁴

3.8 References

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum

Hyperstats Online Statistics Textbook. (n.d.) Retrieved from <http://davidmlane.com/hyperstat/> ¹⁵

Kurtosis. (n.d.). Definition. Retrieved from http://www.statistics.com/index.php?page=glossary&term_id=326 ¹⁶

Kurtosis. (n.d.). *Definition of normality*. Retrieved from <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb> ¹⁷

Onwuegbuzie, A. J., & Daniel, L. G. (2002). Uses and misuses of the correlation coefficient. *Research in the Schools*, 9(1), 73-90.

Skewness. (n.d.) Retrieved from http://www.statistics.com/index.php?page=glossary&term_id=356 ¹⁸

Skewness. (n.d.). *Definition of normality*. Retrieved from <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb> ¹⁹

StatSoft, Inc. (2011). *Electronic statistics textbook*. Tulsa, OK: StatSoft. WEB: <http://www.statsoft.com/textbook/> ²⁰

¹⁴<http://cnx.org/content/col11299/latest/>

¹⁵<http://davidmlane.com/hyperstat/>

¹⁶http://www.statistics.com/index.php?page=glossary&term_id=326

¹⁷<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

¹⁸http://www.statistics.com/index.php?page=glossary&term_id=356

¹⁹<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

²⁰<http://www.statsoft.com/textbook/>

Chapter 4

Calculating Correlations: Parametric and Non Parametric



NOTE: This Chapter has been peer-reviewed, accepted, and endorsed by the National Council of Professors of Educational Administration (NCPEA) as a significant contribution to the scholarship and practice of education administration. Formatted and edited in Connexions by Theodore Creighton and Brad Bizzell, Virginia Tech, Janet Tareilo, Stephen F. Austin State University, and Thomas Kersten, Roosevelt University.

4.1

This chapter is part of a larger Collection (Book) and is available at: *Calculating Basic Statistical Procedures in SPSS: A Self-Help and Practical Guide to Preparing Theses, Dissertations, and Manuscripts* ²

NOTE: Slate and LeBouef have written a "companion book" which is available at: *Preparing and Presenting Your Statistical Findings: Model Write Ups* ³

Authors Information

John R. Slate is a Professor at Sam Houston State University where he teaches Basic and Advanced Statistics courses, as well as professional writing, to doctoral students in Educational Leadership and Counseling. His research interests lie in the use of educational databases, both state and national, to reform school practices. To date, he has chaired and/or served over 100 doctoral student dissertation committees. Recently, Dr. Slate created a website Writing and Statistical Help⁴ to assist students and faculty with both statistical assistance and in editing/writing their dissertations/theses and manuscripts.

¹This content is available online at <<http://cnx.org/content/m37278/1.7/>>.

²<http://my.qoop.com/store/NCPEA-Publications-1781472103076212/>

³<http://my.qoop.com/store/NCPEA-Publications-1781472103076212/>

⁴<http://www.writingandstatisticalhelp.com>

Ana Rojas-LeBouef is a Literacy Specialist at the Reading Center at Sam Houston State University where she teaches developmental reading courses. She recently completed her doctoral degree in Reading, where she conducted a 16-year analysis of Texas statewide data regarding the achievement gap. Her research interests lie in examining the inequities in achievement among ethnic groups. Dr. Rojas-LeBouef also assists students and faculty in their writing and statistical needs on the website Writing and Statistical Help.⁵

Editors Information

Theodore B. Creighton, is a Professor at Virginia Tech and the Publications Director for NCPEA Publications⁶, the Founding Editor of Education Leadership Review,⁷ and the Senior Editor of the NCPEA Connexions Project.

Brad E. Bizzell, is a recent graduate of the Virginia Tech Doctoral Program in Educational Leadership and Policy Studies, and is a School Improvement Coordinator for the Virginia Tech Training and Technical Assistance Center. In addition, Dr. Bizzell serves as an Assistant Editor of the NCPEA Connexions Project in charge of technical formatting and design.

Janet Tareilo, is a Professor at Stephen F. Austin State University and serves as the Assistant Director of NCPEA Publications. Dr. Tareilo also serves as an Assistant Editor of the NCPEA Connexions Project and as a editor and reviewer for several national and international journals in educational leadership.

Thomas Kersten is a Professor at Roosevelt University in Chicago. Dr. Kersten is widely published and an experienced editor and is the author of Taking the Mystery Out of Illinois School Finance⁸, a Connexions Print on Demand publication. He is also serving as Editor in Residence for this book by Slate and LeBouef.

4.2 Calculating Correlations: Parametric and Nonparametric

In this set of steps, readers will calculate either a parametric or a nonparametric statistical analysis, depending on whether the data reflect a normal distribution. A parametric statistical procedure requires that its data be reflective of a normal curve whereas no such assumption is made in the use of a nonparametric procedure. Of the two types of statistical analyses, the parametric procedure is the more powerful one in ascertaining whether or not a statistically significant relationship, in this case, exists. As such, parametric procedures are preferred over nonparametric procedures. When data are not normally distributed, however, parametric analyses may provide misleading and inaccurate results. Accordingly, nonparametric analyses should be used in cases where data are not reflective of a normal curve. In this set of steps, readers are provided with information on how to make the determination of normally or nonnormally distributed data. For detailed information regarding the assumptions underlying parametric and nonparametric procedures, readers are referred to the Hyperstats Online Statistics Textbook at <http://davidmlane.com/hyperstat/>⁹ or to the *Electronic Statistics Textbook* (2011) at <http://www.statsoft.com/textbook/>¹⁰

Research questions for which correlations are appropriate involve asking for relationships between or among variables. The research question, “What is the relationship between study skills and grades for high school students?” could be answered through use of a correlation.

4.2.1 Step One:

Perform ScatterPlots

⁵<http://www.writingandstatisticalhelp.com>

⁶<http://www.ncpeapublications.org>

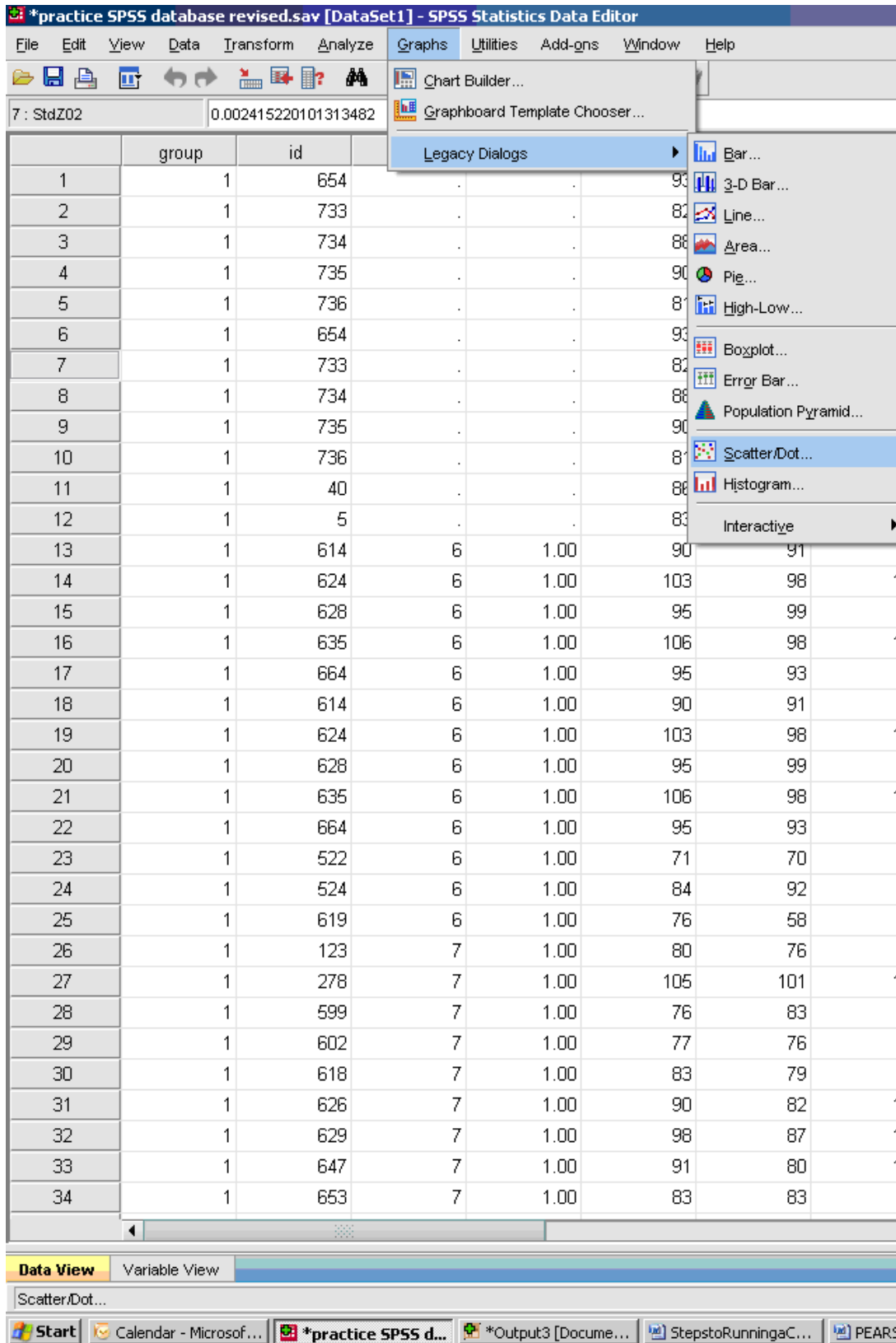
⁷<http://ncpeapublications.org/about-elr.html>

⁸<http://cnx.org/content/col10606/latest/>

⁹<http://davidmlane.com/hyperstat/>

¹⁰<http://www.statsoft.com/textbook/>

- ✓ Graphs
- ✓ Legacy Dialogs
- ✓ Scatter/Dot
- ✓ The Simple Scatter icon should be highlighted



*practice SPSS database revised.sav [DataSet1] - SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

7 : StdZ02 0.002415220101313482

Legacy Dialogs

- Bar...
- 3-D Bar...
- Line...
- Area...
- Pie...
- High-Low...
- Boxplot...
- Error Bar...
- Population Pyramid...
- Scatter/Dot...
- Histogram...
- Interactive

	group	id						
1	1	654	.	.	93			
2	1	733	.	.	82			
3	1	734	.	.	88			
4	1	735	.	.	90			
5	1	736	.	.	81			
6	1	654	.	.	93			
7	1	733	.	.	82			
8	1	734	.	.	88			
9	1	735	.	.	90			
10	1	736	.	.	81			
11	1	40	.	.	88			
12	1	5	.	.	83			
13	1	614	6	1.00	90	91		
14	1	624	6	1.00	103	98	1	
15	1	628	6	1.00	95	99		
16	1	635	6	1.00	106	98	1	
17	1	664	6	1.00	95	93		
18	1	614	6	1.00	90	91		
19	1	624	6	1.00	103	98	1	
20	1	628	6	1.00	95	99		
21	1	635	6	1.00	106	98	1	
22	1	664	6	1.00	95	93		
23	1	522	6	1.00	71	70		
24	1	524	6	1.00	84	92		
25	1	619	6	1.00	76	58		
26	1	123	7	1.00	80	76		
27	1	278	7	1.00	105	101	1	
28	1	599	7	1.00	76	83		
29	1	602	7	1.00	77	76		
30	1	618	7	1.00	83	79		
31	1	626	7	1.00	90	82	1	
32	1	629	7	1.00	98	87	1	
33	1	647	7	1.00	91	80	1	
34	1	653	7	1.00	83	83		

Data View Variable View

Scatter/Dot...

Start Calendar - Microsof... *practice SPSS d... *Output3 [Docume... StepstoRunningaC... PEARS

¹¹<http://cnx.org/content/m37278/latest/figure3.1.PNG/image>

✓ Define

✓ Drag one of the two variables of interest to the first box (Y axis) on the right hand side and the other variable of interest to the second box (X axis) on the right hand side. It does not matter which variable goes in the X or Y axis because your scatterplot results will be the same.

Once you have a variable in each of the two boxes, click on the OK tab on the bottom left hand corner of the screen.

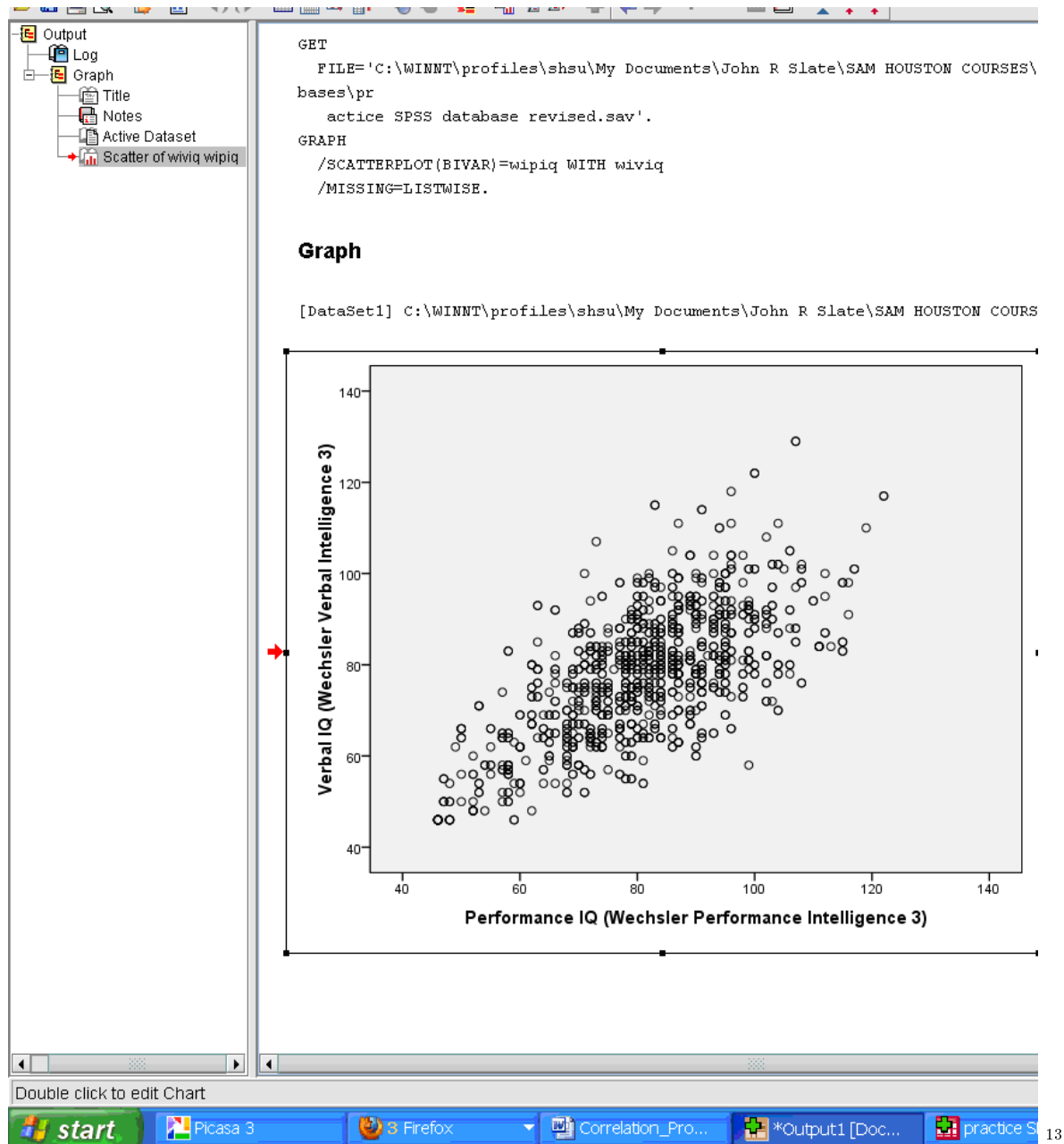
The screenshot shows the SPSS 'Simple Scatterplot' dialog box. The Y-axis is set to 'Verbal IQ (Wechsler Verbal Int...)' and the X-axis is set to 'Disability Group Membership [gr...'. The 'Set Markers by:' and 'Label Cases by:' fields are empty. The 'Panel by:' section has empty 'Rows' and 'Columns' fields. The 'Template' section has 'Use chart specifications from:' set to 'File...'. The background shows a data view table with columns: group, id, age, recodage, wfsiq, wiviq, wipiq, pc, inf, cod. The table contains 34 rows of data.

	group	id	age	recodage	wfsiq	wiviq	wipiq	pc	inf	cod
1	1	654	.	.	93	100	86	8	12	6
2	1	733	.	.	82	78	90	7	4	5
3	1	734	.	.	88	81	99	7	7	7
4	1	735	.	.	88	81	99	7	7	7
5	1	736	.	.	88	81	99	7	7	7
6	1	654	.	.	88	81	99	7	7	7
7	1	733	.	.	88	81	99	7	7	7
8	1	734	.	.	88	81	99	7	7	7
9	1	735	.	.	88	81	99	7	7	7
10	1	736	.	.	88	81	99	7	7	7
11	1	40	.	.	88	81	99	7	7	7
12	1	5	.	.	88	81	99	7	7	7
13	1	614	6	.	88	81	99	7	7	7
14	1	624	6	.	88	81	99	7	7	7
15	1	628	6	.	88	81	99	7	7	7
16	1	635	6	.	88	81	99	7	7	7
17	1	664	6	.	88	81	99	7	7	7
18	1	614	6	.	88	81	99	7	7	7
19	1	624	6	.	88	81	99	7	7	7
20	1	628	6	.	88	81	99	7	7	7
21	1	635	6	.	88	81	99	7	7	7
22	1	664	6	.	88	81	99	7	7	7
23	1	522	6	.	88	81	99	7	7	7
24	1	524	6	.	88	81	99	7	7	7
25	1	619	6	.	88	81	99	7	7	7
26	1	123	7	.	88	81	99	7	7	7
27	1	278	7	.	88	81	99	7	7	7
28	1	599	7	.	88	81	99	7	7	7
29	1	602	7	.	88	81	99	7	7	7
30	1	618	7	1.00	83	79	91	8	5	9
31	1	626	7	1.00	90	82	102	9	10	8
32	1	629	7	1.00	98	87	112	13	8	10
33	1	647	7	1.00	91	80	106	10	4	13
34	1	653	7	1.00	83	83	86	3	7	14

✓ Look at the scatterplots to see whether a linear relationship is present.

In the screenshot below, the relationship is very clearly linear.

¹²<http://cnx.org/content/m37278/latest/figure3.2.PNG/image>



13

4.2.2 Step Two:

Calculate Descriptive Statistics on Variables

- ✓ Analyze
- * Descriptive Statistics
- * Frequencies
- * Click on the variables for which you want descriptive statistics (your dependent variables)
- * You may click on each variable separately or highlight several of them

¹³<http://cnx.org/content/m37278/latest/figure3.3.PNG/image>

* * Once you have a variable in the left hand cell highlighted, click on the arrow in the middle to send the variable to the empty cell titled Variable(s)

The screenshot shows the SPSS Data View window with a dataset containing 34 rows and 10 columns. The columns are: group, id, age, recodage, wifsiq, wiviq, wipiq, pc, inf, and cod. The data is as follows:

	group	id	age	recodage	wifsiq	wiviq	wipiq	pc	inf	cod
1	1	654	.	.	93	100	86	8	12	
2	1	733	.	.	82	78	90	7	4	
3	1	734	.	.	88	81	99	7	7	
4	1	735	.	.	90	91	90	.	.	
5	1	736	.	.	81	78	87	9	6	
6	1	654	.	.	93	100	86	8	12	
7	1	733	.	.	82	78	90	7	4	
8	1	734	.	.	88	81	99	7	7	
9	1	735	.	.	90	91	90	.	.	
10	1	736	.	.	81	78	87	9	6	
11	1	40	.	.						
12	1	5	.	.						
13	1	614	6	1.00						
14	1	624	6	1.00						
15	1	628	6	1.00						
16	1	635	6	1.00						
17	1	664	6	1.00						
18	1	614	6	1.00						
19	1	624	6	1.00						
20	1	628	6	1.00						
21	1	635	6	1.00						
22	1	664	6	1.00						
23	1	522	6	1.00	71	70	78	7	6	
24	1	524	6	1.00	84	92	79	4	5	
25	1	619	6	1.00	76	58	99	9	3	
26	1	123	7	1.00	80	76	87	5	7	
27	1	278	7	1.00	105	101	108	10	9	
28	1	599	7	1.00	76	83	72	5	5	
29	1	602	7	1.00	77	76	81	9	2	
30	1	618	7	1.00	83	79	91	8	5	
31	1	626	7	1.00	90	82	102	9	10	
32	1	629	7	1.00	98	87	112	13	8	
33	1	647	7	1.00	91	80	106	10	4	
34	1	653	7	1.00	83	83	86	3	7	

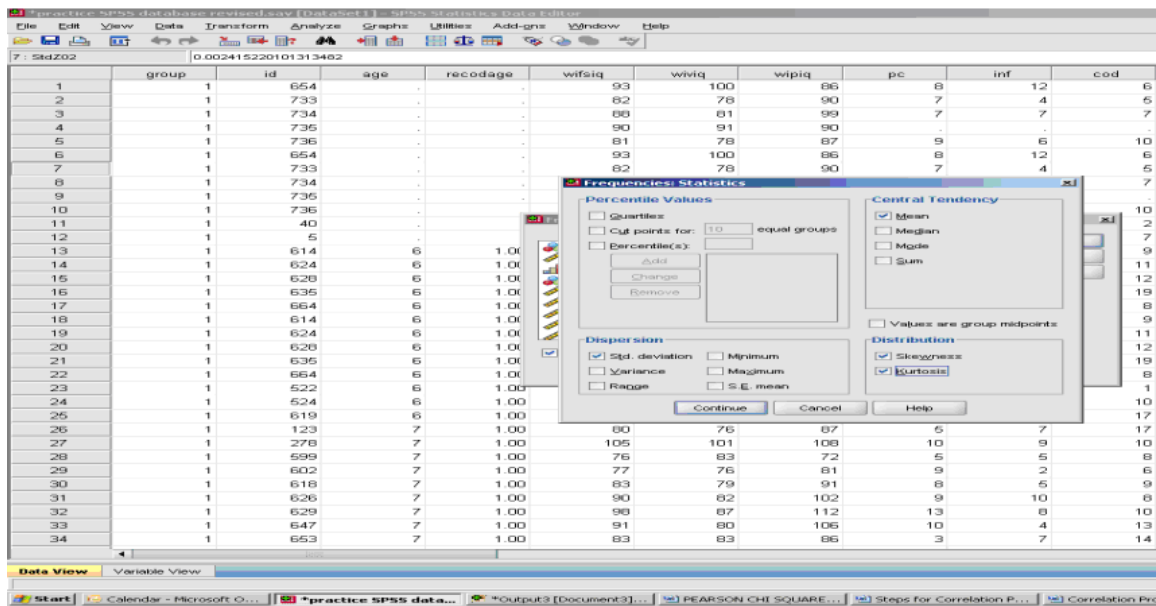
The 'Frequencies' dialog box is open, showing the list of variables on the left and the 'Variable(s):' list on the right. The variable 'Performance IQ (Wechsler Full Scale I...)' is selected in the 'Variable(s):' list. The 'Display frequency tables' checkbox is checked. The dialog has buttons for OK, Paste, Reset, Cancel, and Help.

✓ Statistics

* Click on as many of the options you would like to see results

* At the minimum, click on: *M*, *SD*, Skewness, and Kurtosis

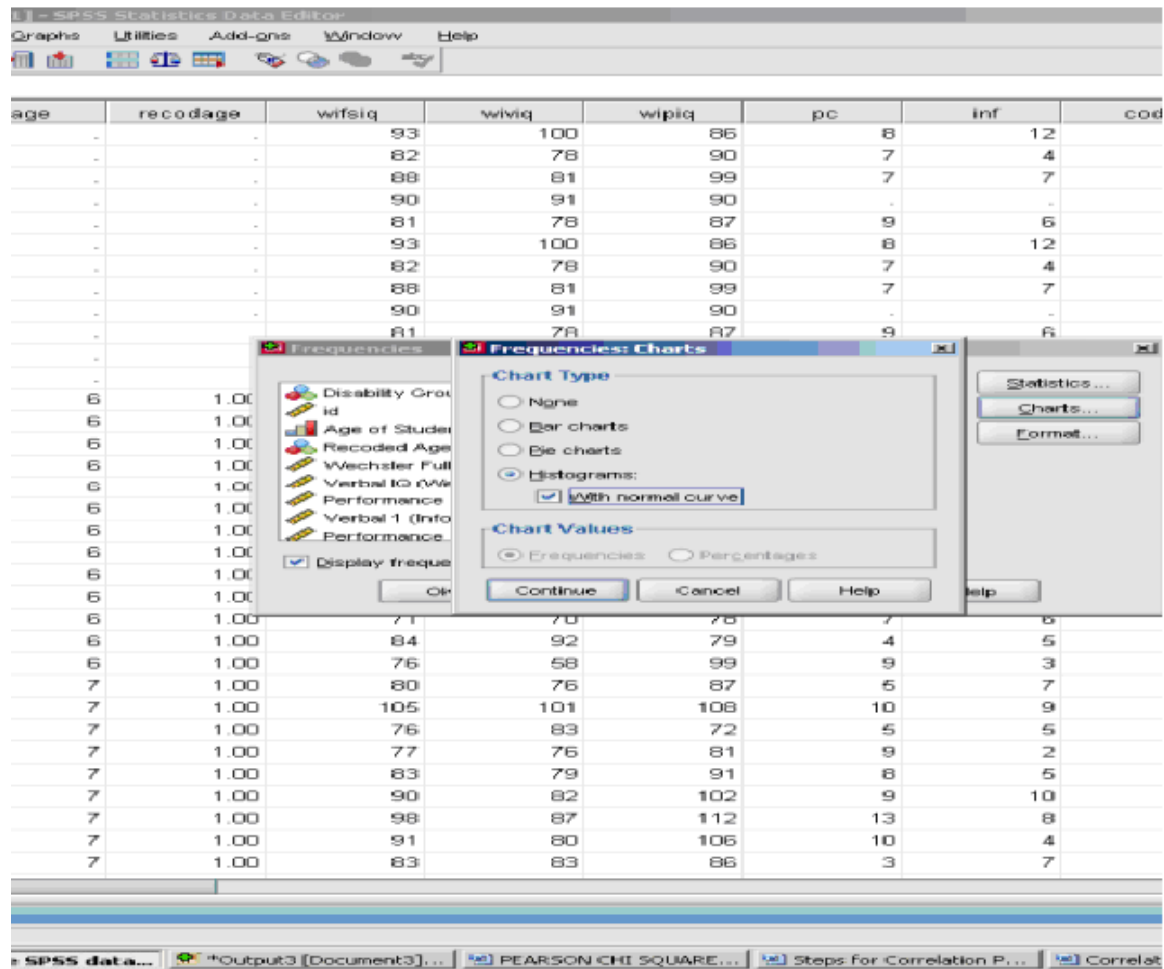
¹⁴<http://cnx.org/content/m37278/latest/figure3.4.PNG/image>



15

- * Continue
- * Charts (these are calculated only if you wish to have visual depictions of skewness and of kurtosis-they are not required)
- * Histograms (not required, optional) with Normal Curve

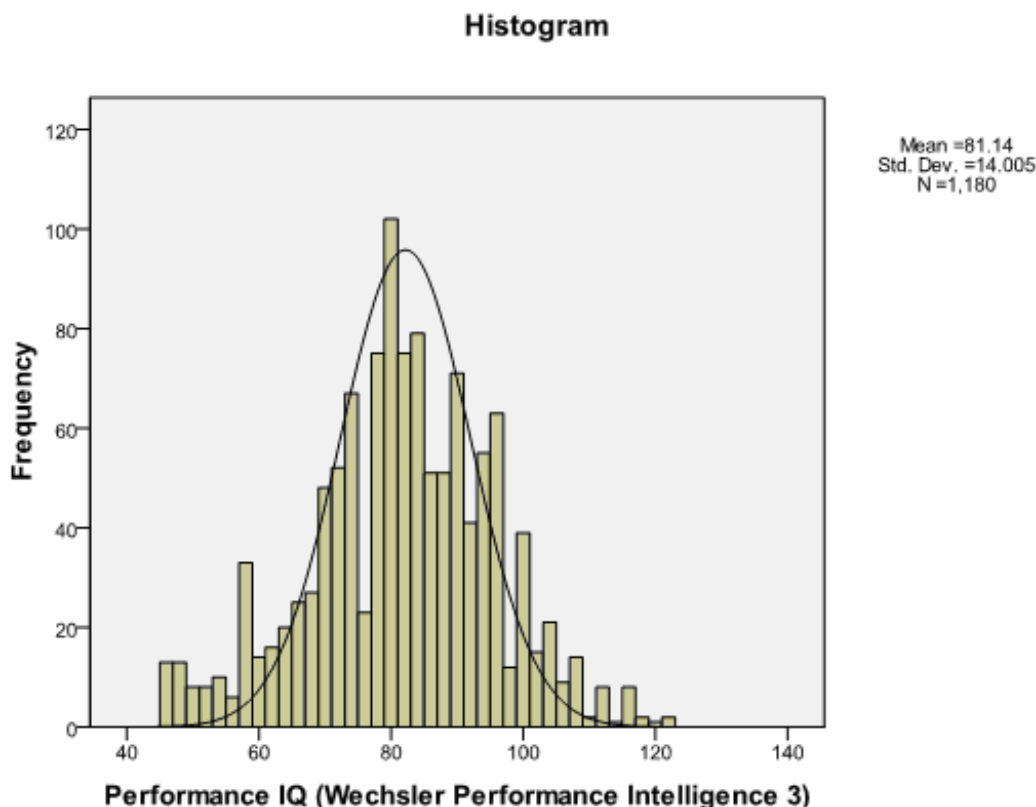
¹⁵<http://cnx.org/content/m37278/latest/figure3.5.PNG/image>



* Continue

* OK

¹⁶<http://cnx.org/content/m37278/latest/figure3.6.PNG/image>



17

4.2.3 Step Three:

Check for Skewness and Kurtosis values falling within/outside the parameters of normality (-3 to +3)

* Skewness [Note. Skewness refers to the extent to which the data are normally distributed around the mean. Skewed data involve having either mostly high scores with a few low ones or having mostly low scores with a few high ones.] Readers are referred to the following sources for a more detailed definition of skewness: http://www.statistics.com/index.php?page=glossary&term_id=356¹⁸ and <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>¹⁹

To standardize the skewness value so that its value can be constant across datasets and across studies, the following calculation must be made: Take the skewness value from the SPSS output (in this case it is -.177) and divide it by the Std. error of skewness (in this case it is .071). If the resulting calculation is within -3 to +3, then the skewness of the dataset is within the range of normality (Onwuegbuzie & Daniel, 2002). If the resulting calculation is outside of this +/-3 range, the dataset is not normally distributed.

* Kurtosis [Note. Kurtosis also refers to the extent to which the data are normally distributed around the mean. This time, the data are piled up higher than normal around the mean or piled up higher than normal at the ends of the distribution.] Readers are referred to the following sources for a more detailed definition of kurtosis: http://www.statistics.com/index.php?page=glossary&term_id=326²⁰ and

¹⁷<http://cnx.org/content/m37278/latest/figure3.7.1.PNG/image>

¹⁸http://www.statistics.com/index.php?page=glossary&term_id=356

¹⁹<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

²⁰http://www.statistics.com/index.php?page=glossary&term_id=326

<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>²¹

To standardize the kurtosis value so that its value can be constant across datasets and across studies, the following calculation must be made: Take the kurtosis value from the SPSS output (in this case it is .072) and divide it by the Std. error of kurtosis (in this case it is .142). If the resulting calculation is within -3 to +3, then the kurtosis of the dataset is within the range of normality (Onwuegbuzie & Daniel, 2002). If the resulting calculation is outside of this +/-3 range, the dataset is not normally distributed.

Performance IQ (Wechsler Performance Intelligence 3)

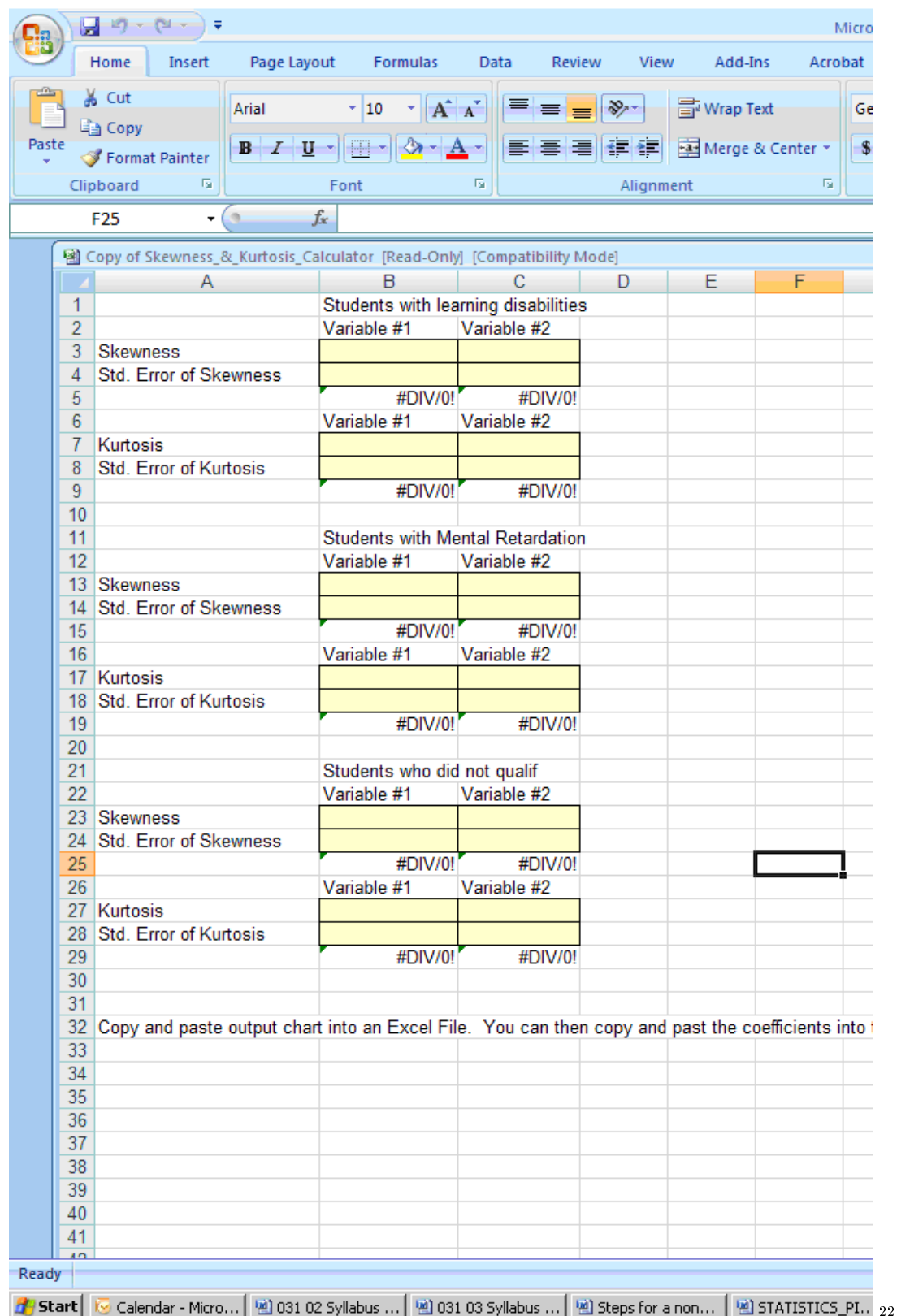
Statistics		
Performance IQ (Wechsler Performance Intelligence 3)		
N	Valid	1180
	Missing	2
Mean		81.14
Std. Deviation		14.005
Skewness		-.177
Std. Error of Skewness		.071
Kurtosis		.072
Std. Error of Kurtosis		.142

Table 4.1

Standardized Coefficients Calculator

Copy variable #1 and #2 into the skewness and kurtosis calculator

²¹<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>



²²<http://cnx.org/content/m37278/latest/figure3.8.PNG/image>

4.2.4 Step Four:

Calculate a Correlation Procedure on the Data

- ✓ Analyze
- ✓ Correlate
- ✓ Bivariate

The screenshot shows the SPSS Statistics Data Editor window with the following data table:

	group	id	wfsiq	wiviq	wipiq
1	1				
2	1				
3	1				
4	1				
5	1				
6	1				
7	1				
8	1				
9	1				
10	1				
11	1				
12	1				
13	1				
14	1				
15	1				
16	1				
17	1				
18	1				
19	1	624	6	1.00	103
20	1	628	6	1.00	95
21	1	635	6	1.00	106
22	1	664	6	1.00	95
23	1	522	6	1.00	71
24	1	524	6	1.00	84
25	1	619	6	1.00	76
26	1	123	7	1.00	80
27	1	278	7	1.00	105
28	1	599	7	1.00	76
29	1	602	7	1.00	77
30	1	618	7	1.00	83
31	1	626	7	1.00	90
32	1	629	7	1.00	98
33	1	647	7	1.00	91
34	1	653	7	1.00	83

The 'Bivariate...' dialog box is open, showing the following options:

- Variables: wfsiq, wiviq, wipiq
- Display: Partial (selected), Bivariate, Distances
- Save: No (selected), Observed, Residuals, Zero
- Options: No (selected), Missing values, Zero

The taskbar at the bottom shows the following open applications: Start, Calendar - Microsoft O..., *practice SPSS data..., *Output3 [Document3]..., and PEARSON CHI SQUARE...

²³<http://cnx.org/content/m37278/latest/figure3.9.PNG/image>

- ✓ Send Over Variables on which you want to calculate a correlation by clicking on the variables in the left hand cell and then clicking on the middle arrow to send them to the right hand cell.
- ✓ Perform a Pearson r if the standardized skewness coefficients and standardized kurtosis coefficients are within normal limits—the Pearson r is the default
- ✓ Calculate a Spearman rho if the standardized skewness coefficients and standardized kurtosis coefficients are outside of the normal limits of ± 3
- ✓ To calculate a Spearman rho, click on the Spearman button and unclick the Pearson r
- ✓ Use the default two-tailed test of significance
- ✓ Use the Flag significant Correlation

The screenshot displays the SPSS interface with a data file named 'StdZ02'. A 'Bivariate Correlations' dialog box is open, showing the selection of variables 'id' and 'wifsiq' for correlation analysis. The dialog box is configured to use Pearson correlation coefficients, a two-tailed test of significance, and to flag significant correlations. The background data table includes columns for group, id, age, recodage, wifsiq, wiviq, wipiq, pc, inf, and cod.

✓ OK

4.2.5 Step Five:

Check for Statistical Significance

1. Go to the correlation box
2. Follow Sig. (2-tailed) row over to chosen variable column
3. If you have any value less than .05 or less than your Bonferroni adjustment, if you are calculating multiple correlations on the same sample in the same study, then you have statistical significance.

²⁴<http://cnx.org/content/m37278/latest/figure3.10.PNG/image>

Correlations

		Verbal IQ (Wechsler Verbal Intelligence 3)	Performance IQ (Wechsler Performance Intelligence 3)
Verbal IQ (Wechsler Verbal Intelligence 3)	Pearson Correlation	1	.664**
	Sig. (2-tailed)		.000
	N	1182	1180
Performance IQ (Wechsler Performance Intelligence 3)	Pearson Correlation	.664**	1
	Sig. (2-tailed)	.000	
	N	1180	1180
**. Correlation is significant at the 0.01 level (2-tailed).			

Table 4.2

NOTE: [In this matrix, it appears that four unique correlations are present, one per cell. In fact, only one unique correlation, or r , is present in this four cell matrix.]

4.2.6 Step Six:

Check For Effect Size

1. Go to the correlation box
2. Find Pearson's Correlation Row or Spearman rho's and follow it to the variable column.
3. Your effect size will be located in the cell where the above intersect.
4. The effect size is calculated as:

4.2.6.1 Cohen's criteria for correlations (1998)

- .1 = small (range from .1 to .29)
- .3 = moderate (range from .3 to .49)
- .5 = large (range from .5 to 1.0)

NOTE: Correlations cannot be greater than 1.00, therefore a 0 should not be placed in front of the decimal.

4.2.7 Step Seven:

Check the Level of Variance the Variables Have in Common

1. Square the Pearson Correlation Value or Spearman rho value to find the variance
2. In this example, the Verbal IQ and the Performance IQ share 44.09% of the variance in common (see correlation value of .664).

Correlations

		Verbal IQ (Wechsler Verbal Intelligence 3)	Performance IQ (Wechsler Performance Intelligence 3)
Verbal IQ (Wechsler Verbal Intelligence 3)	Pearson Correlation	1	.664**
	Sig. (2-tailed)		.000
	N	1182	1180
Performance IQ (Wechsler Performance Intelligence 3)	Pearson Correlation	.664**	1
	Sig. (2-tailed)	.000	
	N	1180	1180
** Correlation is significant at the 0.01 level (2-tailed).			

Table 4.3

4.2.8 Step Eight:

Write the Numerical Sentence

1. $r(n)_{sp=sp}$ correlation coefficient, $_{sp}p_{sp} < .001$ (or Bonferroni-adjusted alpha significance error rate).
2. Using this example: $r(1180) = .66, p < .001$

NOTE: [sp means to insert a space.] Remember that all mathematical symbols are placed in italics.

4.2.9 Step Nine:

Narrative and Interpretation

1. r value
2. sample size or n
3. p value
4. r^2 value
5. $r(1180) = .66, p < .001$, 44.09% of variance accounted for.
6. Note that the r value itself is the effect size.

4.3 Writing Up Your Statistics

So, how do you "write up" your Research Questions and your Results? Schuler W. Huck (2000) in his seminal book entitled, *Reading Statistics and Research*, points to the importance of your audience understanding and making sense of your research in written form. Huck further states:

4.3.1

This book is designed to help people *decipher* what researchers are trying to communicate in the written or oral summaries of their investigations. Here, the goal is simply to distill meaning from the words, symbols, tables, and figures included in the research report. To be competent in this arena, one must not only be able to decipher what's presented but also to "fill in the holes"; this is the case because researchers typically assume that those receiving the research report are familiar with unmentioned details of the research process and statistical treatment of data.

Writing Up Your Correlations

Researchers and Professors John Slate and Ana Rojas-LeBouef understand this critical issue, so often neglected or not addressed by other authors and researchers. They point to the importance of doctoral students "writing up their statistics" in a way that others can understand your reporting and as importantly, interpret the meaning of your significant findings and implications for the preparation and practice of educational leadership. Slate and LeBouef provide you with a model for "writing up your Parametric and Non-Parametric Correlations statistics."

Click here to view: *Writing Up Your Parametric Correlation Statistics* ²⁵

Click here to view: *Writing Up Your Nonparametric Correlation Statistics* ²⁶

4.4 References

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum

Hyperstats Online Statistics Textbook. (n.d.) Retrieved from <http://davidmlane.com/hyperstat/> ²⁷

Kurtosis. (n.d.). Definition. Retrieved from http://www.statistics.com/index.php?page=glossary&term_id=326 ²⁸

Kurtosis. (n.d.). *Definition of normality*. Retrieved from <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb> ²⁹

Onwuegbuzie, A. J., & Daniel, L. G. (2002). Uses and misuses of the correlation coefficient. *Research in the Schools*, 9(1), 73-90.

Skewness. (n.d.) Retrieved from http://www.statistics.com/index.php?page=glossary&term_id=356 ³⁰

Skewness. (n.d.). *Definition of normality*. Retrieved from <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb> ³¹

StatSoft, Inc. (2011). *Electronic statistics textbook*. Tulsa, OK: StatSoft. WEB: <http://www.statsoft.com/textbook/> ³²

²⁵<http://cnx.org/content/col11299/latest/>

²⁶<http://cnx.org/content/col11299/latest/>

²⁷<http://davidmlane.com/hyperstat/>

²⁸http://www.statistics.com/index.php?page=glossary&term_id=326

²⁹<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

³⁰http://www.statistics.com/index.php?page=glossary&term_id=356

³¹<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

³²<http://www.statsoft.com/textbook/>

Chapter 5

Conducting a Parametric Independent Samples t-test



NOTE: This chapter has been peer-reviewed, accepted, and endorsed by the National Council of Professors of Educational Administration (NCPEA) as a significant contribution to the scholarship and practice of education administration. Formatted and edited in Connexions by Theodore Creighton and Brad Bizzell, Virginia Tech, Janet Tareilo, Stephen F. Austin State University, and Thomas Kersten, Roosevelt University.

5.1

This chapter is part of a larger Collection (Book) and is available at: *Calculating Basic Statistical Procedures in SPSS: A Self-Help and Practical Guide to Preparing Theses, Dissertations, and Manuscripts* ²

NOTE: Slate and LeBouef have written a "companion book" which is available at: *Preparing and Presenting Your Statistical Findings: Model Write Ups* ³

Authors Information

John R. Slate is a Professor at Sam Houston State University where he teaches Basic and Advanced Statistics courses, as well as professional writing, to doctoral students in Educational Leadership and Counseling. His research interests lie in the use of educational databases, both state and national, to reform school practices. To date, he has chaired and/or served over 100 doctoral student dissertation committees. Recently, Dr. Slate created a website Writing and Statistical Help⁴ to assist students and faculty with both statistical assistance and in editing/writing their dissertations/theses and manuscripts.

¹This content is available online at <<http://cnx.org/content/m37279/1.5/>>.

²<http://my.qoop.com/store/NCPEA-Publications-1781472103076212/>

³<http://my.qoop.com/store/NCPEA-Publications-1781472103076212/>

⁴<http://www.writingandstatisticalhelp.com>

Ana Rojas-LeBouef is a Literacy Specialist at the Reading Center at Sam Houston State University where she teaches developmental reading courses. She recently completed her doctoral degree in Reading, where she conducted a 16-year analysis of Texas statewide data regarding the achievement gap. Her research interests lie in examining the inequities in achievement among ethnic groups. Dr. Rojas-LeBouef also assists students and faculty in their writing and statistical needs on the website Writing and Statistical Help.⁵

Editors Information

Theodore B. Creighton, is a Professor at Virginia Tech and the Publications Director for NCPEA Publications⁶, the Founding Editor of Education Leadership Review,⁷ and the Senior Editor of the NCPEA Connexions Project.

Brad E. Bizzell, is a recent graduate of the Virginia Tech Doctoral Program in Educational Leadership and Policy Studies, and is a School Improvement Coordinator for the Virginia Tech Training and Technical Assistance Center. In addition, Dr. Bizzell serves as an Assistant Editor of the NCPEA Connexions Project in charge of technical formatting and design.

Janet Tareilo, is a Professor at Stephen F. Austin State University and serves as the Assistant Director of NCPEA Publications. Dr. Tareilo also serves as an Assistant Editor of the NCPEA Connexions Project and as a editor and reviewer for several national and international journals in educational leadership.

Thomas Kersten is a Professor at Roosevelt University in Chicago. Dr. Kersten is widely published and an experienced editor and is the author of Taking the Mystery Out of Illinois School Finance⁸, a Connexions Print on Demand publication. He is also serving as Editor in Residence for this book by Slate and LeBouef.

5.2 Conducting a Parametric Independent Samples t-test

In this set of steps, readers will calculate either a parametric or a nonparametric statistical analysis, depending on whether the data for the dependent variable reflect a normal distribution. A parametric statistical procedure requires that its data be reflective of a normal curve whereas no such assumption is made in the use of a nonparametric procedure. Of the two types of statistical analyses, the parametric procedure is the more powerful one in ascertaining whether or not a statistically significant difference, in this case, exists. As such, parametric procedures are preferred over nonparametric procedures. When data are not normally distributed, however, parametric analyses may provide misleading and inaccurate results. According, nonparametric analyses should be used in cases where data are not reflective of a normal curve. In this set of steps, readers are provided with information on how to make the determination of normally or nonnormally distributed data. For detailed information regarding the assumptions underlying parametric and nonparametric procedures, readers are referred to the Hyperstats Online Statistics Textbook at <http://davidmlane.com/hyperstat/>⁹ or to the *Electronic Statistics Textbook* (2011) at <http://www.statsoft.com/textbook/>¹⁰

For this parametric independent samples *t*-test to be appropriately used, at least half of the standardized skewness coefficients and the standardized kurtosis coefficients must be within the normal range (± 3 , Onwuegbuzie & Daniel, 2002). Research questions for which independent samples *t*-tests are appropriate involve asking for differences in a dependent variable by group membership (i.e., only two groups are present for *t*-tests). The research question, “What is the difference between boys and girls in their science performance among middle school students?” could be answered through use of an independent samples *t*-test.

⁵<http://www.writingandstatisticalhelp.com>

⁶<http://www.ncpeapublications.org>

⁷<http://ncpeapublications.org/about-elr.html>

⁸<http://cnx.org/content/col10606/latest/>

⁹<http://davidmlane.com/hyperstat/>

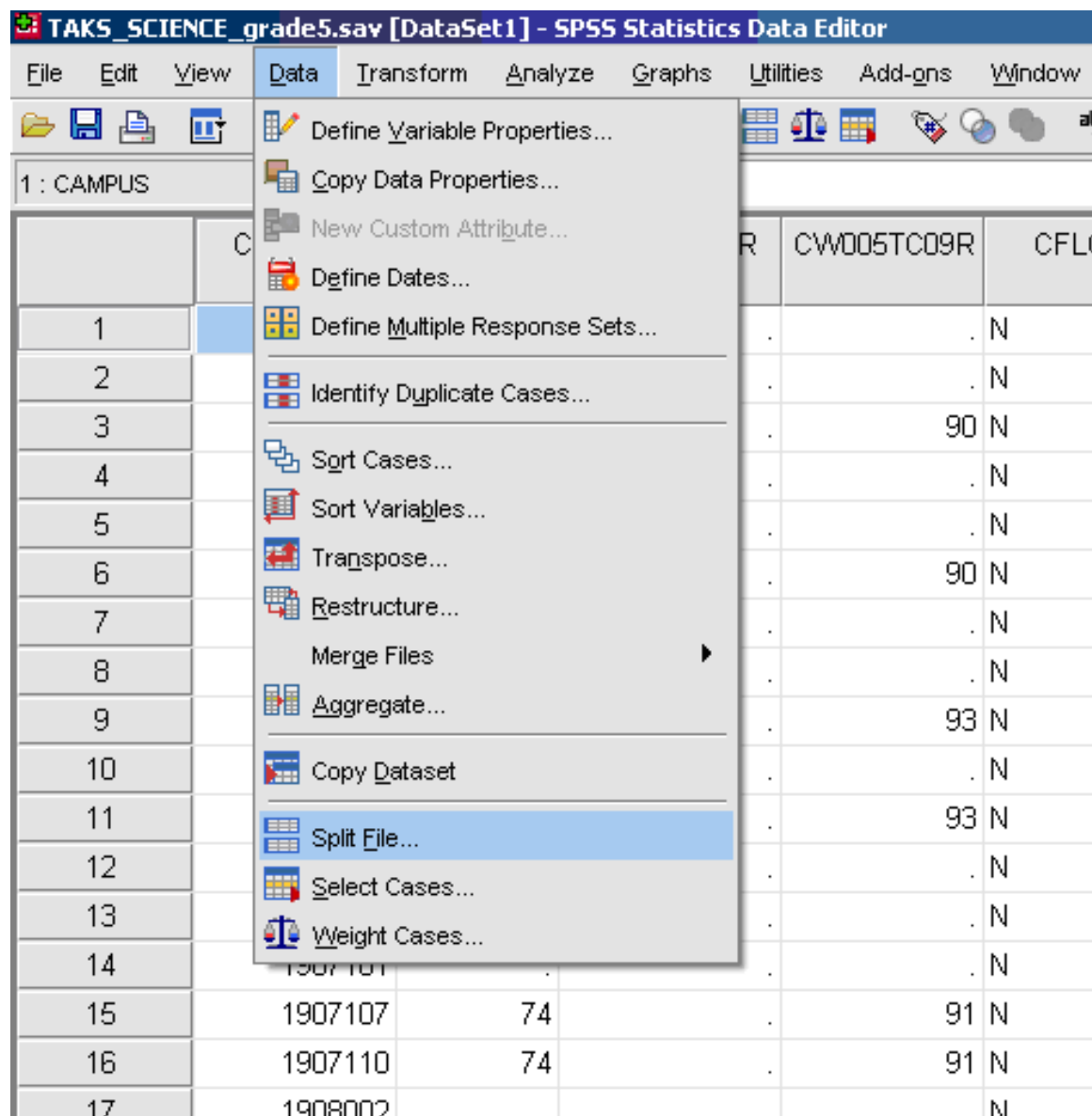
¹⁰<http://www.statsoft.com/textbook/>

5.2.1 Step One

Calculate Frequencies on the Split Groups

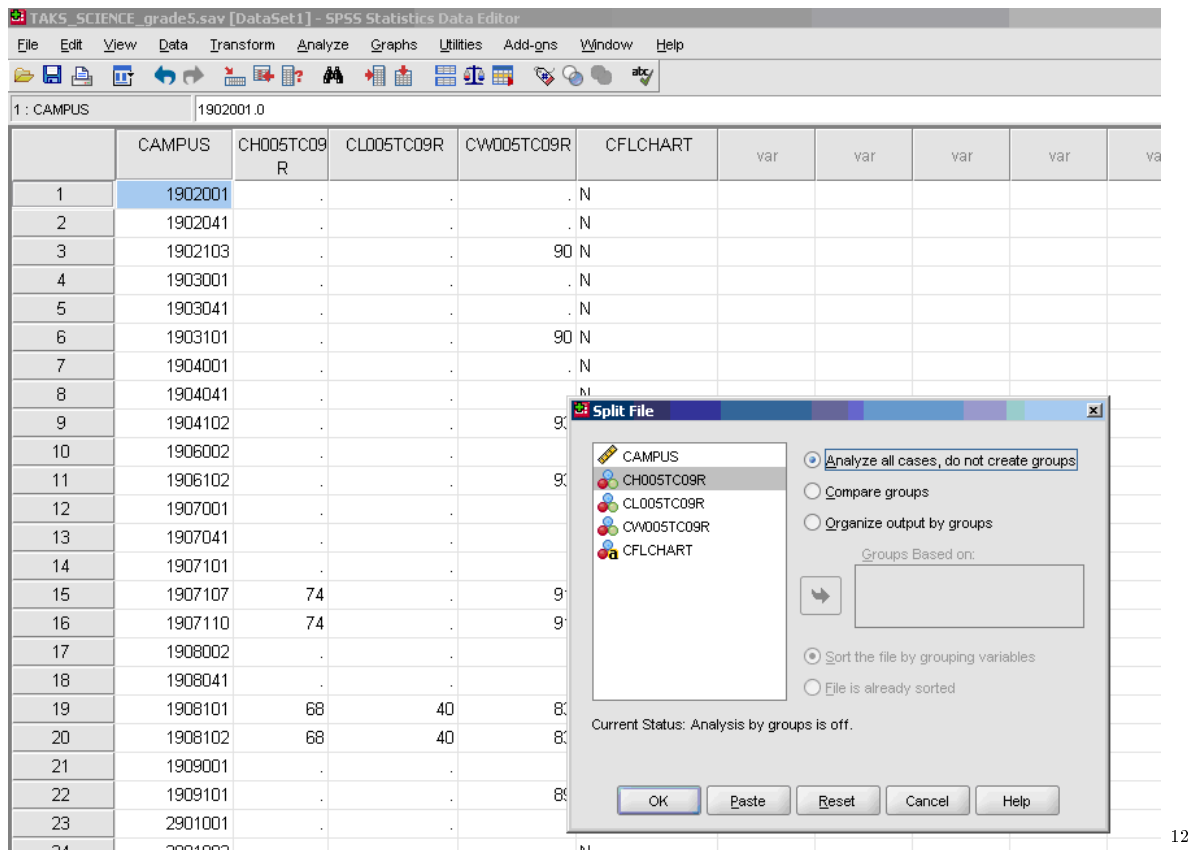
✓ Data

* Split File



Your screen will show that all cases are going to be analyzed and a “do not create groups”. You will need to click the compare groups and move the independent variable over to the “Group Based on”. In the case of a *t*-test, the grouping variable or independent variable will consist of two groups.

¹¹<http://cnx.org/content/m37279/latest/figure4.1.PNG/image>



After you do this, your screen should resemble the following:

¹²<http://cnx.org/content/m37279/latest/figure4.2.PNG/image>

SPSS Statistics Data Editor window showing a dataset with columns: group, id, age, recodage, wifsiq, wiviq, wipiq, pc, inf, co. The 'Split File' dialog box is open, showing 'id' as the grouping variable. The 'Compare groups' radio button is selected. The 'Groups Based on:' field contains 'Disability Group Membership [...]'. The 'Sort the file by grouping variables' radio button is selected. The 'Current Status: Analysis by groups is off.' message is displayed at the bottom of the dialog.

13

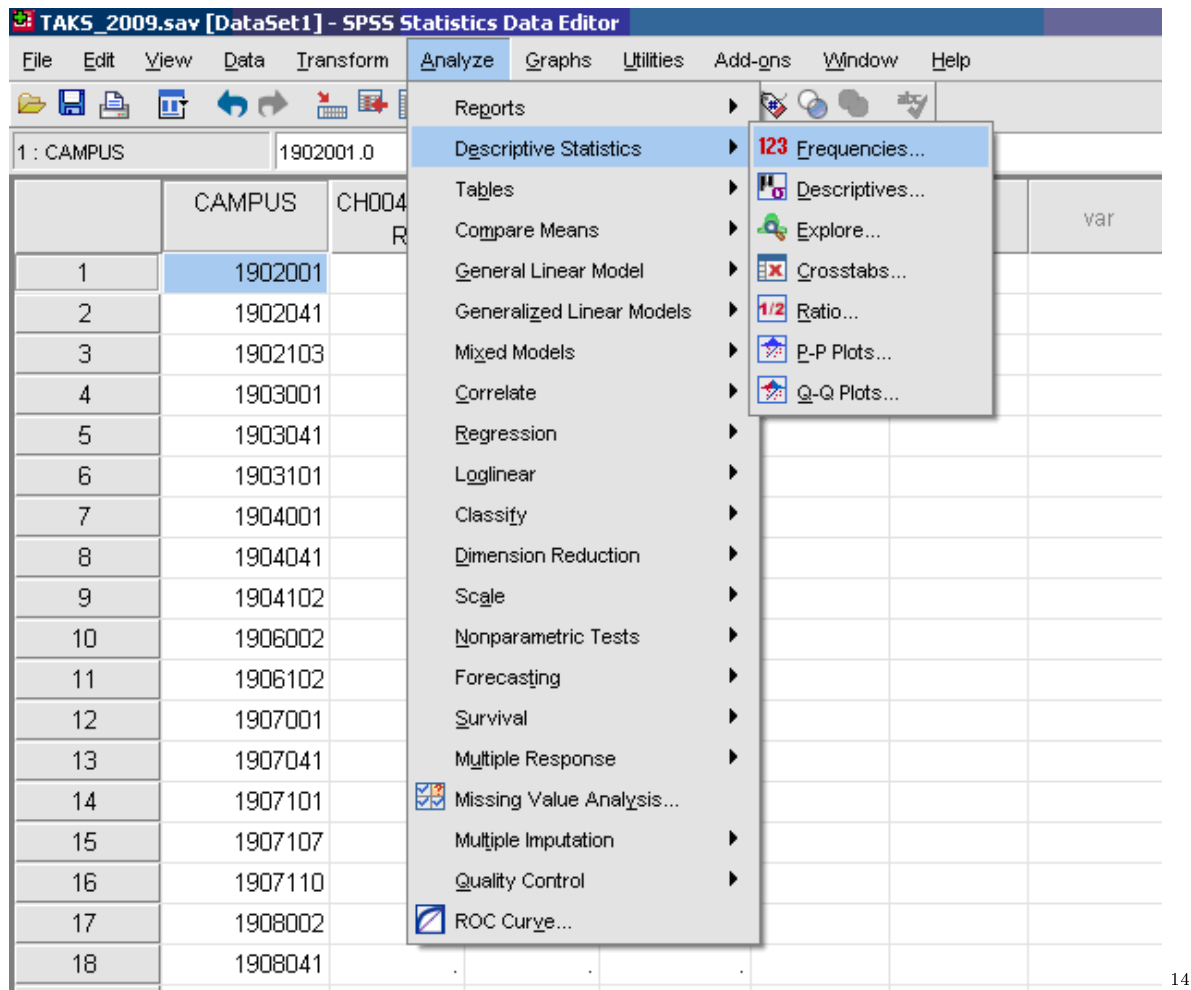
* Then click OK

✓ Analyze

* Descriptive Statistics

* Frequencies

¹³<http://cnx.org/content/m37279/latest/figure4.3.PNG/image>



✓ Move over the dependent (outcome) variable

¹⁴<http://cnx.org/content/m37279/latest/figure4.4.PNG/image>

The screenshot shows the SPSS Statistics Data Editor window with a dataset named 'TAKS_SCIENCE_grade5.sav [DataSet1]'. The data is organized into rows and columns. The columns are labeled CAMPUS, CH005TC09R, CL005TC09R, CW005TC09R, and CFLCHART. The rows are numbered 1 through 23. A 'Frequencies' dialog box is open, showing the variable 'CH005TC09R' selected for analysis. The dialog box includes options for 'Statistics...', 'Charts...', 'Format...', and 'Display frequency tables'.

15

✓ Statistics

* Mean

* Standard Deviation

* Skewness [Note. Skewness refers to the extent to which the data are normally distributed around the mean. Skewed data involve having either mostly high scores with a few low ones or having mostly low scores with a few high ones.] Readers are referred to the following sources for a more detailed definition of skewness: http://www.statistics.com/index.php?page=glossary&term_id=356¹⁶ and <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>¹⁷

To standardize the skewness value so that its value can be constant across datasets and across studies, the following calculation must be made: Take the skewness value from the SPSS output and divide it by the Std. error of skewness. If the resulting calculation is within -3 to +3, then the skewness of the dataset is within the range of normality (Onwuegbuzie & Daniel, 2002). If the resulting calculation is outside of this +/-3 range, the dataset is not normally distributed.

* Kurtosis [Note. Kurtosis also refers to the extent to which the data are normally distributed around the mean. This time, the data are piled up higher than normal around the mean or piled up higher than normal at the ends of the distribution.] Readers are referred to the following sources for a more detailed definition of kurtosis: http://www.statistics.com/index.php?page=glossary&term_id=326¹⁸ and <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>¹⁹

To standardize the kurtosis value so that its value can be constant across datasets and across studies, the following calculation must be made: Take the kurtosis value from the SPSS output and divide it by the Std. error of kurtosis. If the resulting calculation is within -3 to +3, then the kurtosis of the dataset

¹⁵<http://cnx.org/content/m37279/latest/figure4.5.PNG/image>

¹⁶http://www.statistics.com/index.php?page=glossary&term_id=356

¹⁷<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

¹⁸http://www.statistics.com/index.php?page=glossary&term_id=326

¹⁹<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

is within the range of normality (Onwuegbuzie & Daniel, 2002). If the resulting calculation is outside of this ± 3 range, the dataset is not normally distributed.

* Continue

* OK

The screenshot shows the SPSS Statistics Data Editor window with the file 'TAKS_SCIENCE_grade5.sav [DataSet1]'. The data table has 26 rows and 11 columns. The first column is 'CAMPUS' with values ranging from 1902001 to 3902004. The next four columns are 'CH005TC09R', 'CL005TC09R', 'CW005TC09R', and 'CFLCHART'. The last five columns are labeled 'var'. A 'Frequencies: Statistics' dialog box is open, showing the following options:

- Percentile Values:**
 - ☐ Quartiles
 - ☐ Cut points for: 10 equal groups
 - ☐ Percentile(s):
 - Buttons: Add, Change, Remove
- Central Tendency:**
 - ☒ Mean
 - ☐ Median
 - ☐ Mode
 - ☐ Sum
 - ☐ Values are group midpoints
- Dispersion:**
 - ☒ Std. deviation
 - ☐ Variance
 - ☐ Range
 - ☐ Minimum
 - ☐ Maximum
 - ☐ S.E. mean
- Distribution:**
 - ☒ Skewness
 - ☒ Kurtosis

The 'Continue' button is highlighted in the dialog box.

✓ Charts (these are calculated only if you wish to have visual depictions of skewness and of kurtosis-they are not required)

* Histogram~ with normal curve (not required, optional)

✓ Continue

✓ OK

²⁰<http://cnx.org/content/m37279/latest/figure4.6.PNG/image>

23 :

	CAMPUS	CH004TR09 R	CL004TR09 R	CW004TR09 R	var	var	var	var	var	var
1	1902001	.	.	.						
2	1902041	.	.	.						
3	1902103	.	.	97						
4	1903001	.	.	.						
5	1903041	.	.	.						
6	1903101	89	.	97						
7	1904001	.	.	.						
8	1904041	.	.	.						
9	1904102	.	.	87						
10	1906002	.	.	.						
11	1906102	.	.	88						
12	1907001	.	.	.						
13	1907041	.	.	.						
14	1907101	.	.	.						
15	1907107	82	33	91						
16	1907110	82	33	91						
17	1908002	.	.	.						
18	1908041	.	.	.						
19	1908101	57	.	90						
20	1908102	57	.	90						
21	1909001	.	.	.						
22	1909101	.	.	70						
23	2901001	.	.	.						
24	2901002	.	.	.						

21

NOTE: Before you continue to another application you must complete the following:

- ✓ Data
- ✓ Split Files
- ✓ Analyze all cases, do not create groups
- ✓ OK

²¹ <http://cnx.org/content/m37279/latest/figure4.7.PNG/image>

7 : StdZ02 0.002415220101313482

	group	id	age	recodage	wifsiq	wiviq	wipiq	pc	inf	co
1	1	654	.	.	93	100	86	8	12	
2	1	733	.	.	82	78	90	7	4	
3	1	734	.	.	88	81	99	7	7	
4	1	735	.	.	90	91	90	.	.	
5	1	736	.	.	81	78	87	9	6	
6	1	654	.	.	93	100	86	8	12	
7	1	733	.	.	82	78	90	7	4	
8	1	734	.	.	88	81	99	7	7	
9	1	735	.	.						
10	1	736	.	.						
11	1	40	.	.						
12	1	5	.	.						
13	1	614	6	1.00						
14	1	624	6	1.00						
15	1	628	6	1.00						
16	1	635	6	1.00						
17	1	664	6	1.00						
18	1	614	6	1.00						
19	1	624	6	1.00						
20	1	628	6	1.00						
21	1	635	6	1.00						
22	1	664	6	1.00						
23	1	522	6	1.00						
24	1	524	6	1.00						
25	1	619	6	1.00	76	58	99	9	3	
26	1	123	7	1.00	80	76	87	5	7	

5.3

5.3.1 Step Two

Check for Skewness and Kurtosis values falling within/without the parameters of normality (-3 to +3).

Note that each variable below has its own skewness and its own kurtosis values. Thus, a total of three standardized skewness coefficients and three standardized kurtosis coefficients can be calculated from information in the table below.

		CH005TC09R	CL005TC09R	CW005TC09R
N	Valid	3125	1805	1877
	Missing	5197	6517	6445

continued on next page

²²<http://cnx.org/content/m37279/latest/figure4.8.PNG/image>

Skewness	-1.129	-.479	-2.197
Std. Error of Skewness	.044	.058	.056
Kurtosis	1.818	-.412	6.991
Std. Error of Kurtosis	.088	.115	.113

Table 5.1: Skewness and Kurtosis Coefficients

Standardized Coefficients Calculator

Copy variable #1 and #2 into the skewness and kurtosis calculator

The screenshot shows a Microsoft Excel spreadsheet titled 'Copy of Skewness & Kurtosis Calculator [Read-Only] [Compatibility Mode]'. The spreadsheet is organized into three main sections, each corresponding to a different group of students. Each section contains rows for Skewness, Std. Error of Skewness, Kurtosis, and Std. Error of Kurtosis, with columns for Variable #1 and Variable #2. The results for Variable #1 and Variable #2 are mostly '#DIV/0!' errors. A text box at the bottom right says 'Copy and paste output chart into an Excel File. You can then copy and paste the coefficients into the appropriate blocks'.

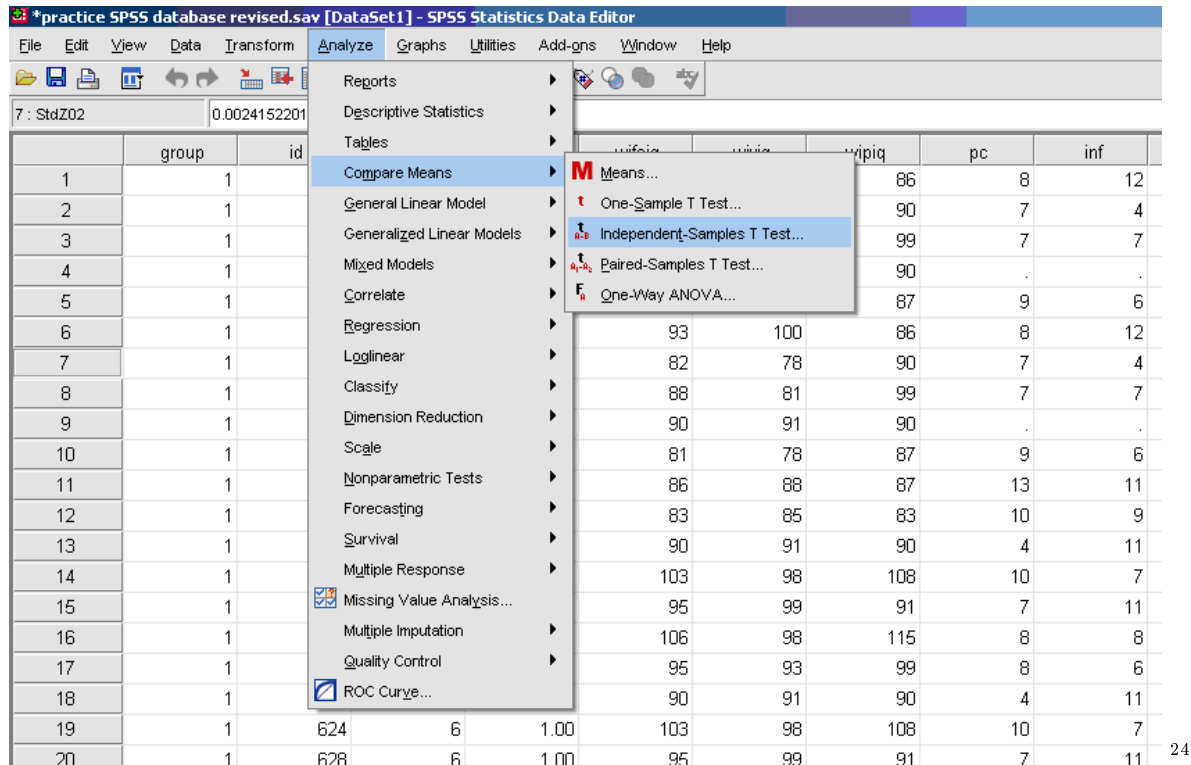
Note. Prior to calculating parametric independent t -tests, at least half of your standardized coefficients should be within the ± 3 range.

5.4 Step Three

Calculate a Parametric Independent Samples t -test on Data (after you have unsplit your file)

²³<http://cnx.org/content/m37279/latest/figure4.9.PNG/image>

- ✓ Analyze
- ✓ Compare Means
- ✓ Independent Samples t -test



- ✓ Test Variable would be your Dependent Variable (e.g., test scores)
- ✓ Grouping Variable would be your dichotomous Independent Variable

²⁴<http://cnx.org/content/m37279/latest/figure4.10.PNG/image>

7 : StdZ02 0.002415220101313482

	group	id	age	recodage	wifsiq	wiviq	wipiq	pc	inf	cod
1	1	654	.	.	93	100	86	8	12	6
2	1	733	.	.	82	78	90	7	4	5
3	1	734	.	.	88	81	99	7	7	7
4	1	735	.	.	90	91	90	.	.	.
5	1	736	.	.	81	78	87	9	6	10
6	1	654	.	.	93	100	86	8	12	6
7	1	733	.	.	82	78	90	7	4	5
8	1	734	.	.	88	81	99	7	7	7
9	1	735	.	.	90	91	90	.	.	.
10	1	736	.	.	81	78	87	9	6	10
11	1	40	.	.	.	.	.	.	.	.
12	1	5	.	.	.	.	.	.	.	.
13	1	614	6	1.00	.	.	.	.	.	.
14	1	624	6	1.00	.	.	.	.	.	.
15	1	628	6	1.00	.	.	.	.	.	.
16	1	635	6	1.00	.	.	.	.	.	.
17	1	664	6	1.00	.	.	.	.	.	.
18	1	614	6	1.00	.	.	.	.	.	.
19	1	624	6	1.00	.	.	.	.	.	.
20	1	628	6	1.00	.	.	.	.	.	.
21	1	635	6	1.00	.	.	.	.	.	.
22	1	664	6	1.00	.	.	.	.	.	.
23	1	522	6	1.00	71	70	78	7	6	1
24	1	524	6	1.00	84	92	79	4	5	10
25	1	619	6	1.00	76	58	99	9	3	17
26	1	123	7	1.00	80	76	87	5	7	17
27	1	778	7	1.00	105	101	102	10	9	10

Independent-Samples T Test

Test Variable(s): Verbal IQ (Wechsler Ve...

Grouping Variable: group(? ?)

Define Groups...

✓ Define Groups

✓ Group One is No. 1 and Group Two is No. 2 (or whatever numbers you used to identify each group)

Note: Click on view than value labels to find the code for each group.

✓ Continue

7 : StdZ02 0.002415220101313482

	group	id	age	recodage	wifsiq	wiviq	wipiq	pc	inf	cod
1	1	654	.	.	93	100	86	8	12	6
2	1	733	.	.	82	78	90	7	4	5
3	1	734	.	.	88	81	99	7	7	7
4	1	735	.	.	90	91	90	.	.	.
5	1	736	.	.	81	78	87	9	6	10
6	1	654	.	.	93	100	86	8	12	6
7	1	733	.	.	82	78	90	7	4	5
8	1	734	.	.	88	81	99	7	7	7
9	1	735	.	.	90	91	90	.	.	.
10	1	736	.	.	81	78	87	9	6	10
11	1	40	.	.	.	.	.	.	.	.
12	1	5	.	.	.	.	.	.	.	.
13	1	614	6	1.00	.	.	.	.	.	.
14	1	624	6	1.00	.	.	.	.	.	.
15	1	628	6	1.00	.	.	.	.	.	.
16	1	635	6	1.00	.	.	.	.	.	.
17	1	664	6	1.00	.	.	.	.	.	.
18	1	614	6	1.00	.	.	.	.	.	.
19	1	624	6	1.00	.	.	.	.	.	.
20	1	628	6	1.00	.	.	.	.	.	.
21	1	635	6	1.00	.	.	.	.	.	.
22	1	664	6	1.00	.	.	.	.	.	.
23	1	522	6	1.00	71	70	78	7	6	1
24	1	524	6	1.00	84	92	79	4	5	10

Independent-Samples T Test

Test Variable(s):

Define Groups

Use specified values

Group 1: 1

Group 2: 2

Continue

²⁵<http://cnx.org/content/m37279/latest/figure4.11.PNG/image>

²⁶<http://cnx.org/content/m37279/latest/figure4.12.PNG/image>

5.5 Step Four

Check for Statistical Significance

- * Go to the Independent Samples Test Box (bottom row ~ Equal variances not assumed) and look at the cell labeled Sig. (2-tailed) to check for significance. Always use the bottom row.
- * If you have any value less than .05 then you have statistical significance, unless you have adjusted for multiple statistical analyses using the Bonferroni procedure. Remember to replace the third zero with a 1, if the sig value is .000 (i.e., if the sig value reads as .000, replace the third 0, so it reads as .001). If you calculate more than one *t*-test, you must use the

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	T	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Verbal IQ (Wechsler Verbal Intelligence 3)	Equal variances assumed	15.880	.000	32.258	775	.000	22.946	.711	21.550	24.342
	Equal variances not assumed			34.672	686.950	.000	22.946	.662	21.647	24.245

Table 5.2: Independent Samples Test

1. Numerical Sentence = $t(df)_{sp} = t_{sp} p_{sp} < .001$ (or Bonferroni-adjusted alpha significance error rate).
 - df is located in Independent Samples Box
 - t is located in Independent Samples Box
2. Numerical sentence is written as: $t(686.95) = 34.67$ $p < .001$, example was statistically significant.

5.6 Writing Up Your Statistics

So, how do you "write up" your Research Questions and your Results? Schuler W. Huck (2000) in his seminal book entitled, *Reading Statistics and Research*, points to the importance of your audience understanding and making sense of your research in written form. Huck further states:

5.6.1

This book is designed to help people *decipher* what researchers are trying to communicate in the written or oral summaries of their investigations. Here, the goal is simply to distill meaning from the words, symbols,

tables, and figures included in the research report. To be competent in this arena, one must not only be able to decipher what's presented but also to "fill in the holes"; this is the case because researchers typically assume that those receiving the research report are familiar with unmentioned details of the research process and statistical treatment of data.

Researchers and Professors John Slate and Ana Rojas-LeBouef understand this critical issue, so often neglected or not addressed by other authors and researchers. They point to the importance of doctoral students "writing up their statistics" in a way that others can understand your reporting and as importantly, interpret the meaning of your significant findings and implications for the preparation and practice of educational leadership. Slate and LeBouef provide you with a model for "writing up your Independent Samples t -test statistics."

Click here to view: *Writing Up Your Independent Samples t -test Statistics* ²⁷

5.7 References

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum

Hyperstats Online Statistics Textbook. (n.d.) Retrieved from <http://davidmlane.com/hyperstat/> ²⁸

Kurtosis. (n.d.). Definition. Retrieved from http://www.statistics.com/index.php?page=glossary&term_id=326 ²⁹

Kurtosis. (n.d.). *Definition of normality*. Retrieved from <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb> ³⁰

Onwuegbuzie, A. J., & Daniel, L. G. (2002). Uses and misuses of the correlation coefficient. *Research in the Schools*, 9(1), 73-90.

Skewness. (n.d.) Retrieved from http://www.statistics.com/index.php?page=glossary&term_id=356 ³¹

Skewness. (n.d.). *Definition of normality*. Retrieved from <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb> ³²

StatSoft, Inc. (2011). *Electronic statistics textbook*. Tulsa, OK: StatSoft. WEB: <http://www.statsoft.com/textbook/> ³³

²⁷<http://cnx.org/content/col11299/latest/>

²⁸<http://davidmlane.com/hyperstat/>

²⁹http://www.statistics.com/index.php?page=glossary&term_id=326

³⁰<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

³¹http://www.statistics.com/index.php?page=glossary&term_id=356

³²<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

³³<http://www.statsoft.com/textbook/>

Chapter 6

Conducting a Parametric Dependent Samples t-test (Paired Samples t-test)



NOTE: This chapter has been peer-reviewed, accepted, and endorsed by the National Council of Professors of Educational Administration (NCPEA) as a significant contribution to the scholarship and practice of education administration. Formatted and edited in Connexions by Theodore Creighton and Brad Bizzell, Virginia Tech, Janet Tareilo, Stephen F. Austin State University, and Thomas Kersten, Roosevelt University.

6.1

This chapter is part of a larger Collection (Book) and is available at: *Calculating Basic Statistical Procedures in SPSS: A Self-Help and Practical Guide to Preparing Theses, Dissertations, and Manuscripts* ²

NOTE: Slate and LeBouef have written a "companion book" which is available at: *Preparing and Presenting Your Statistical Findings: Model Write Ups* ³

Authors Information

John R. Slate is a Professor at Sam Houston State University where he teaches Basic and Advanced Statistics courses, as well as professional writing, to doctoral students in Educational Leadership and Counseling. His research interests lie in the use of educational databases, both state and national, to reform school practices. To date, he has chaired and/or served over 100 doctoral student dissertation committees. Recently, Dr. Slate created a website Writing and Statistical Help ⁴ to assist students and faculty with both statistical assistance and in editing/writing their dissertations/theses and manuscripts.

¹This content is available online at <<http://cnx.org/content/m37328/1.6/>>.

²<http://my.qoop.com/store/NCPEA-Publications-1781472103076212/>

³<http://my.qoop.com/store/NCPEA-Publications-1781472103076212/>

⁴<http://www.writingandstatisticalhelp.com>

Ana Rojas-LeBouef is a Literacy Specialist at the Reading Center at Sam Houston State University where she teaches developmental reading courses. She recently completed her doctoral degree in Reading, where she conducted a 16-year analysis of Texas statewide data regarding the achievement gap. Her research interests lie in examining the inequities in achievement among ethnic groups. Dr. Rojas-LeBouef also assists students and faculty in their writing and statistical needs on the website Writing and Statistical Help.⁵

Editors Information

Theodore B. Creighton, is a Professor at Virginia Tech and the Publications Director for NCPEA Publications⁶, the Founding Editor of Education Leadership Review,⁷ and the Senior Editor of the NCPEA Connexions Project.

Brad E. Bizzell, is a recent graduate of the Virginia Tech Doctoral Program in Educational Leadership and Policy Studies, and is a School Improvement Coordinator for the Virginia Tech Training and Technical Assistance Center. In addition, Dr. Bizzell serves as an Assistant Editor of the NCPEA Connexions Project in charge of technical formatting and design.

Janet Tareilo, is a Professor at Stephen F. Austin State University and serves as the Assistant Director of NCPEA Publications. Dr. Tareilo also serves as an Assistant Editor of the NCPEA Connexions Project and as a editor and reviewer for several national and international journals in educational leadership.

Thomas Kersten is a Professor at Roosevelt University in Chicago. Dr. Kersten is widely published and an experienced editor and is the author of Taking the Mystery Out of Illinois School Finance⁸, a Connexions Print on Demand publication. He is also serving as Editor in Residence for this book by Slate and LeBouef.

6.2 Conducting a Parametric Dependent Samples t -test

In this set of steps, readers will calculate either a parametric or a nonparametric statistical analysis, depending on whether the data for the dependent variable reflect a normal distribution. A parametric statistical procedure requires that its data be reflective of a normal curve whereas no such assumption is made in the use of a nonparametric procedure. Of the two types of statistical analyses, the parametric procedure is the more powerful one in ascertaining whether or not a statistically significant difference, in this case, exists. As such, parametric procedures are preferred over nonparametric procedures. When data are not normally distributed, however, parametric analyses may provide misleading and inaccurate results. According, nonparametric analyses should be used in cases where data are not reflective of a normal curve. In this set of steps, readers are provided with information on how to make the determination of normally or nonnormally distributed data. For detailed information regarding the assumptions underlying parametric and nonparametric procedures, readers are referred to the Hyperstats Online Statistics Textbook at <http://davidmlane.com/hyperstat/>⁹ or to the *Electronic Statistics Textbook* (2011) at <http://www.statsoft.com/textbook/>¹⁰

For this parametric dependent samples t -test to be appropriately used, at least half of the standardized skewness coefficients and the standardized kurtosis coefficients must be within the normal range ($+/-3$, Onwuegbuzie & Daniel, 2002). Research questions for which dependent samples t -tests are appropriate involve asking for differences in a dependent variable by group membership (i.e., only two groups are present for t -tests and, in this case, must be connected). The research question, “What is the effect of a reading intervention program on science performance among elementary school students?” could be answered through use of an dependent samples t -test.

⁵<http://www.writingandstatisticalhelp.com>

⁶<http://www.ncpeapublications.org>

⁷<http://ncpeapublications.org/about-elr.html>

⁸<http://cnx.org/content/col10606/latest/>

⁹<http://davidmlane.com/hyperstat/>

¹⁰<http://www.statsoft.com/textbook/>

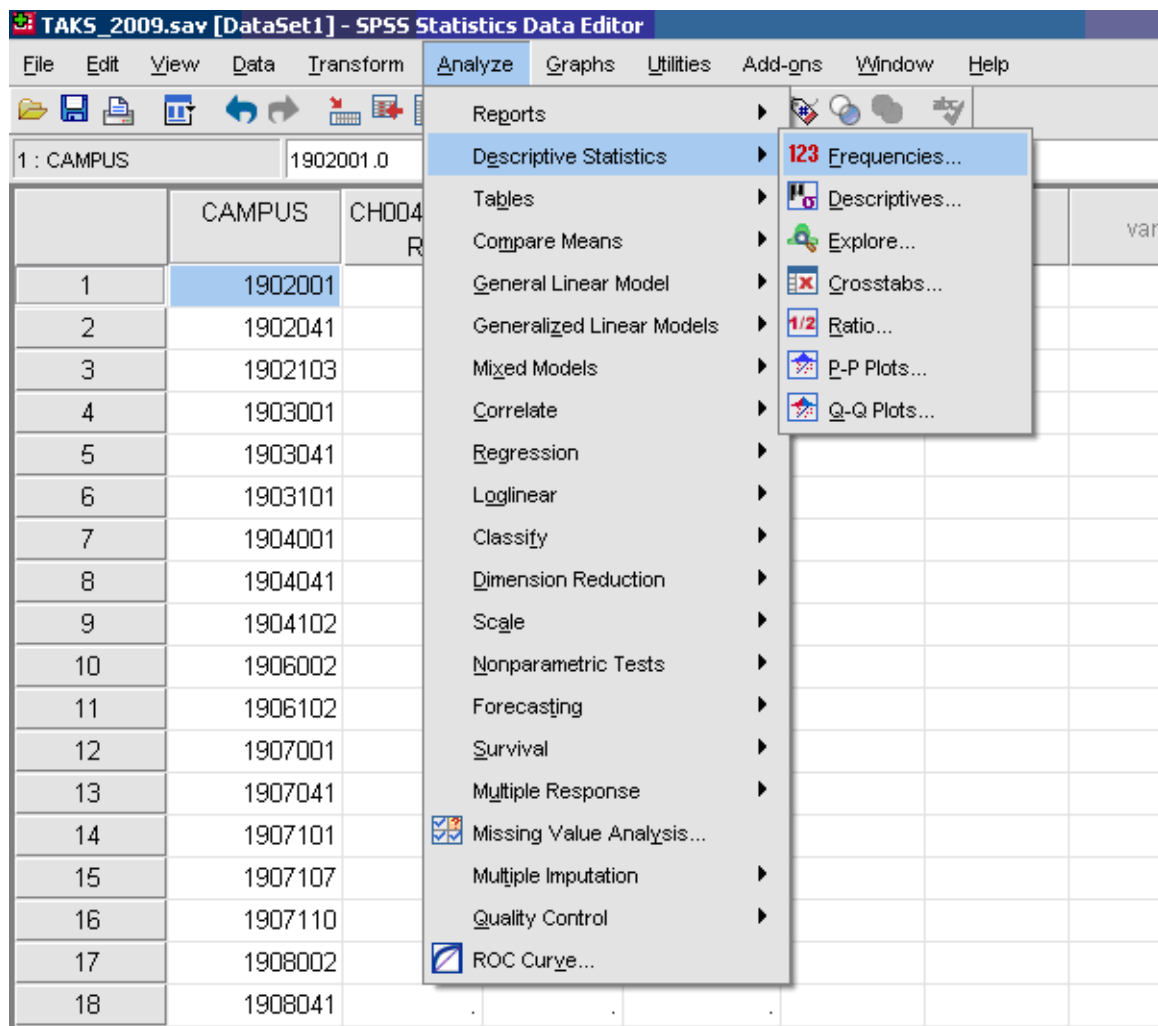
6.3 Step One:

Compute Measures of Normality for the Dependent Variable

✓ Analyze

* Descriptive Statistics

* Frequencies



✓ Move over the dependent (outcome) variable

¹¹<http://cnx.org/content/m37328/latest/figure5.1.PNG/image>

The screenshot shows the SPSS Statistics Data Editor window with a dataset named 'TAKS_SCIENCE_grade5.sav [DataSet1]'. The dataset has 23 rows and 9 columns. The first column is 'CAMPUS', followed by 'CH005TC09R', 'CL005TC09R', 'CW005TC09R', 'CFLCHART', and three empty columns labeled 'var'. The 'Frequencies' dialog box is open, showing 'CAMPUS' selected in the left list and 'CH005TC09R' selected in the 'Variable(s):' list on the right. The dialog also has buttons for Statistics..., Charts..., and Format..., and checkboxes for 'Display frequency tables'.

12

✓ Statistics

* **Skewness** [Note. Skewness refers to the extent to which the data are normally distributed around the mean. Skewed data involve having either mostly high scores with a few low ones or having mostly low scores with a few high ones.] Readers are referred to the following sources for a more detailed definition of skewness: http://www.statistics.com/index.php?page=glossary&term_id=356¹³ and <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>¹⁴

To standardize the skewness value so that its value can be constant across datasets and across studies, the following calculation must be made: Take the skewness value from the SPSS output and divide it by the Std. error of skewness. If the resulting calculation is within -3 to +3, then the skewness of the dataset is within the range of normality (Onwuegbuzie & Daniel, 2002). If the resulting calculation is outside of this +/-3 range, the dataset is not normally distributed.

* **Kurtosis** [Note. Kurtosis also refers to the extent to which the data are normally distributed around the mean. This time, the data are piled up higher than normal around the mean or piled up higher than normal at the ends of the distribution.] Readers are referred to the following sources for a more detailed definition of kurtosis: http://www.statistics.com/index.php?page=glossary&term_id=326¹⁵ and <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>¹⁶

To standardize the kurtosis value so that its value can be constant across datasets and across studies, the following calculation must be made: Take the kurtosis value from the SPSS output and divide it by the Std. error of kurtosis. If the resulting calculation is within -3 to +3, then the kurtosis of the dataset is within the range of normality (Onwuegbuzie & Daniel, 2002). If the resulting calculation is outside

¹²<http://cnx.org/content/m37328/latest/figure5.2.PNG/image>

¹³http://www.statistics.com/index.php?page=glossary&term_id=356

¹⁴<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

¹⁵http://www.statistics.com/index.php?page=glossary&term_id=326

¹⁶<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

of this ± 3 range, the dataset is not normally distributed.

* Continue

* OK

The screenshot shows the SPSS Statistics Data Editor window with a dataset named 'TAKS_2009.sav'. The data is organized into columns: CAMPUS, CH004TR09R, CL004TR09R, and CW004TR09R. A 'Frequencies: Statistics' dialog box is open, displaying various statistical options. In the 'Distribution' section, 'Skewness' and 'Kurtosis' are checked, while 'Std. deviation', 'Variance', 'Range', 'Minimum', 'Maximum', and 'S.E. mean' are unchecked. The 'Central Tendency' section has 'Mean', 'Median', 'Mode', and 'Sum' all unchecked. The 'Percentile Values' section has 'Quartiles', 'Cut points for: 10 equal groups', and 'Percentile(s):' all unchecked. The 'Dispersion' section has 'Std. deviation', 'Variance', 'Range', 'Minimum', 'Maximum', and 'S.E. mean' all unchecked. The 'Values are group midpoints' checkbox is also unchecked. The 'Continue', 'Cancel', and 'Help' buttons are at the bottom of the dialog box.

17

* Uncheck the "display frequency tables" so that you are not provided with the frequencies of your data every time descriptive statistics are obtained.

6.4 Step Two:

Check for Skewness and Kurtosis values falling within/without the parameters of normality (-3 to +3). Note that each variable below has its own skewness and its own kurtosis values. Thus, a total of three standardized skewness coefficients and three standardized kurtosis coefficients can be calculated from information in the table below.

		CH005TC09R	CL005TC09R	CW005TC09R
<i>continued on next page</i>				

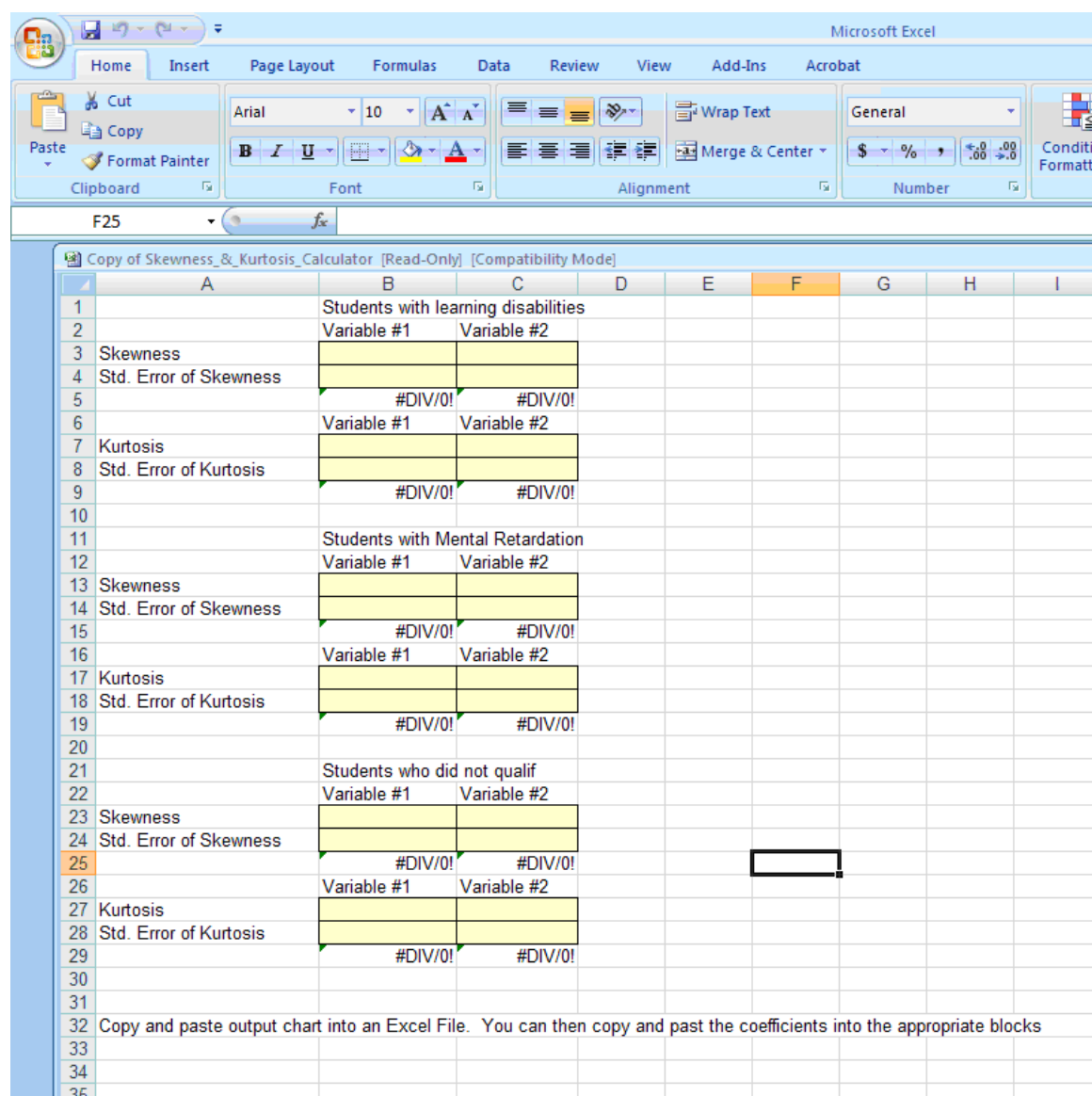
¹⁷<http://cnx.org/content/m37328/latest/figure5.3.PNG/image>

N	Valid	3125	1805	1877
	Missing	5197	6517	6445
Skewness		-1.129	-.479	-2.197
Std. Error of Skewness		.044	.058	.056
Kurtosis		1.818	-.412	6.991
Std. Error of Kurtosis		.088	.115	.113

Table 6.1: Skewness and Kurtosis Coefficients

Standardized Coefficients Calculator

Copy variable #1 and #2 into the skewness and kurtosis calculator

¹⁸<http://cnx.org/content/m37328/latest/figure5.4.PNG/image>

✓ Charts (these are calculated only if you wish to have visual depictions of skewness and of kurtosis-they are not required)

* Histogram~ with normal curve (not required, optional)

The screenshot shows the SPSS Statistics Data Editor window with a dataset named 'TAKS_2009.sav [DataSet1]'. The data is organized into columns: CAMPUS, CH004TR09, CL004TR09, and CW004TR09. The 'Frequencies: Charts' dialog box is open, showing the 'Chart Type' section with 'None' selected. The 'Chart Values' section has 'Frequencies' selected. The 'With normal curve' checkbox is unchecked. The 'Display frequency' checkbox is checked. The 'Continue' button is highlighted.

	CAMPUS	CH004TR09	CL004TR09	CW004TR09
1	1902001	.	.	.
2	1902041	.	.	.
3	1902103	.	.	97
4	1903001	.	.	.
5	1903041	.	.	.
6	1903101	89	.	97
7	1904001	.	.	.
8	1904041	.	.	.
9	1904102	.	.	87
10	1906002	.	.	.
11	1906102	.	.	88
12	1907001	.	.	.
13	1907041	.	.	.
14	1907101	.	.	.
15	1907107	82	33	91
16	1907110	82	33	91
17	1908002	.	.	.
18	1908041	.	.	.
19	1908101	57	.	90
20	1908102	57	.	90
21	1909001	.	.	.
22	1909101	.	.	70
23	2901001	.	.	.

19

6.5 Step Three:

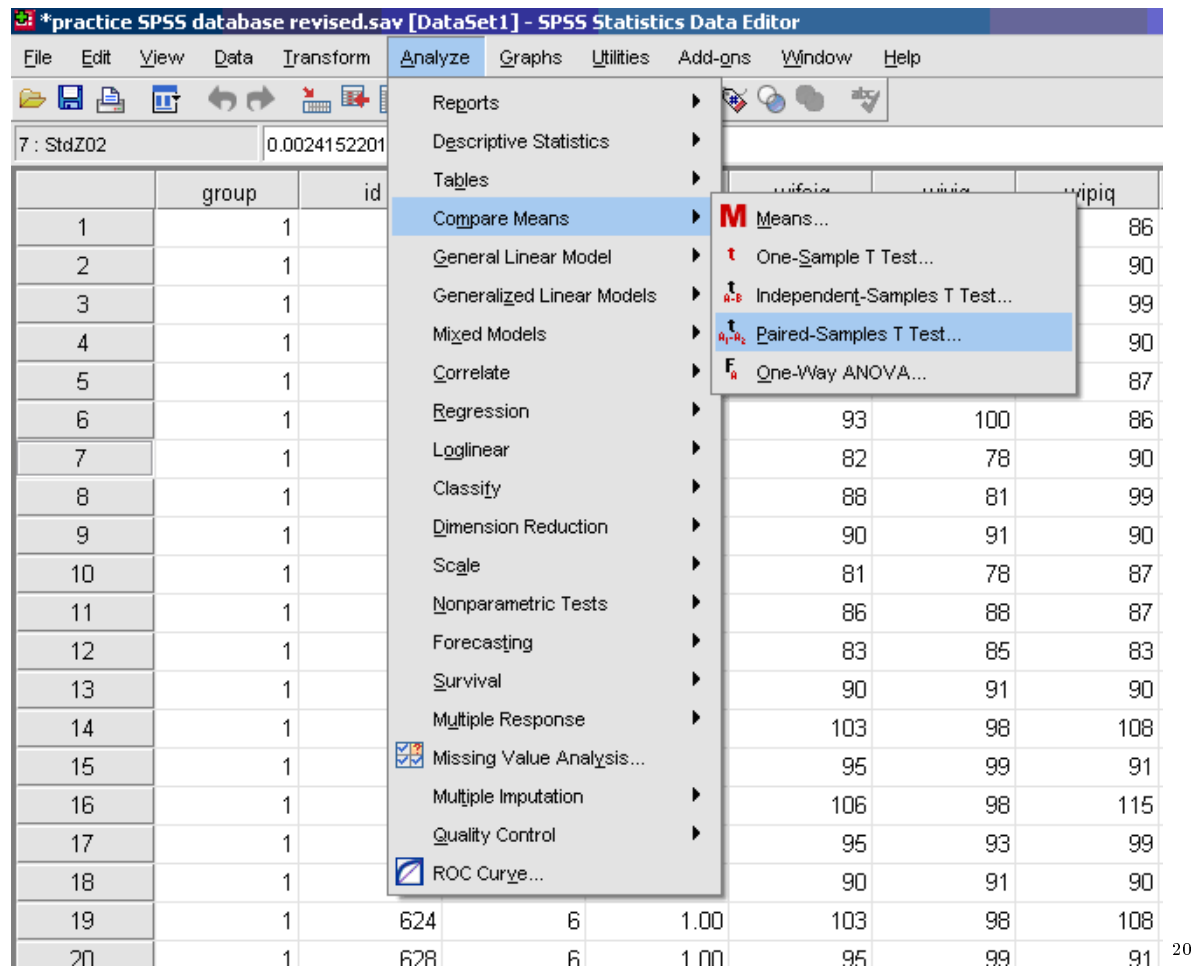
Calculate Paired Samples t -test on Data

✓ Analyze

✓ Compare Means

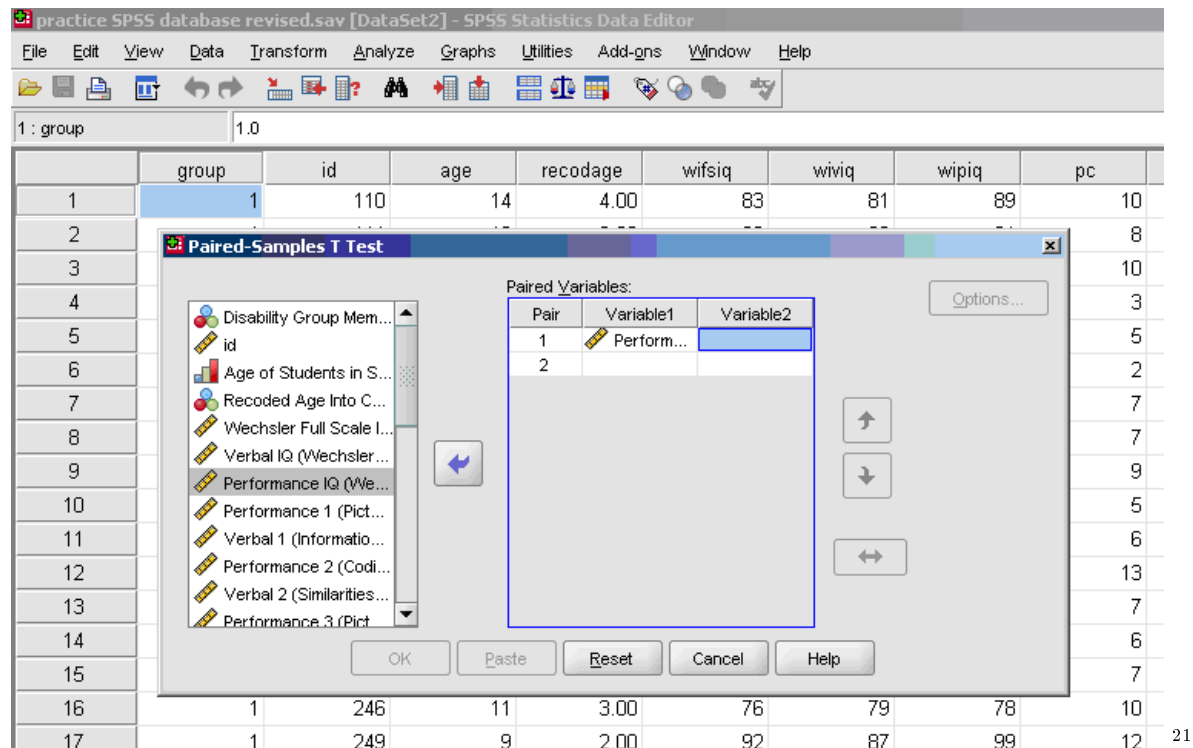
✓ Paired samples t -test

¹⁹<http://cnx.org/content/m37328/latest/figure5.5.PNG/image>



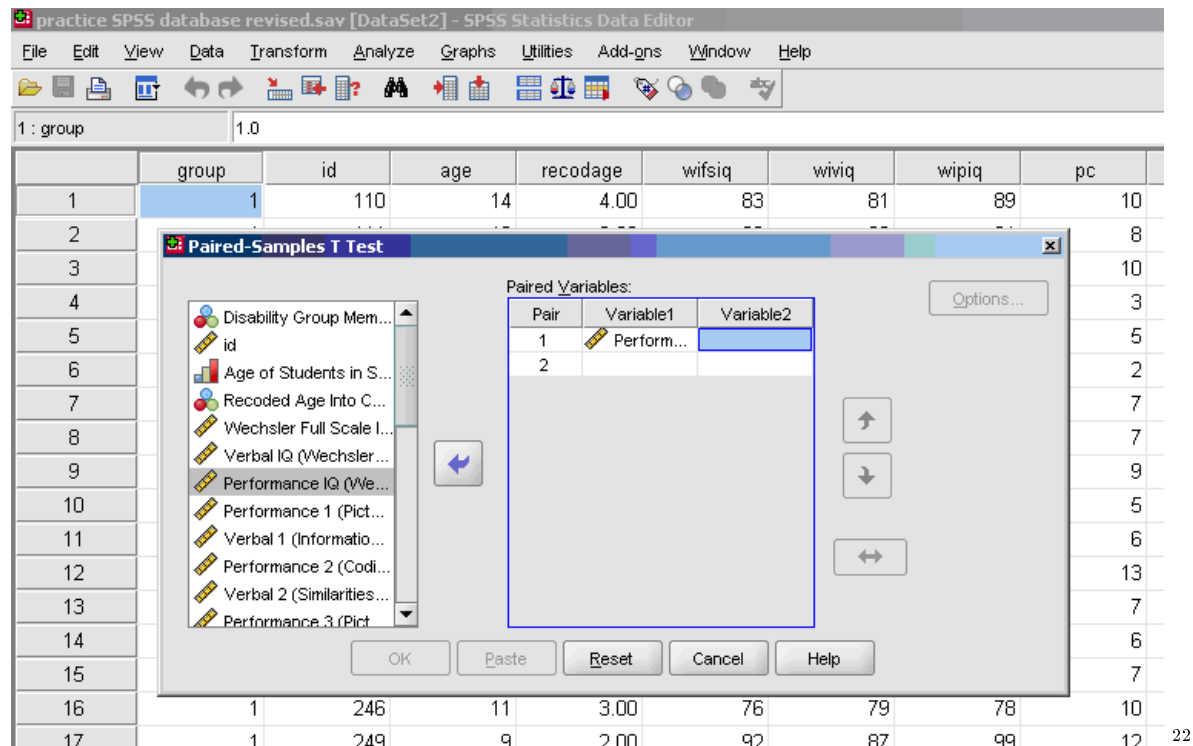
- ✓ Click on one dependent variable
- ✓ Arrow to send over to Paired Variables Side, Variable 1

²⁰<http://cnx.org/content/m37328/latest/figure5.6.PNG/image>



- ✓ Click on second dependent variable
- ✓ Arrow to send over to Paired Variables Side, Variable 2

²¹ <http://cnx.org/content/m37328/latest/figure5.7.1.PNG/image>



✓ OK

6.6 Step Four:

Check for Statistical Significance

Go to the Paired Samples Test Box and look at the very last cell labeled Sig. (2-tailed) to check for significance.

If you have any value less than .05 then you have statistical significance. Remember to replace the third zero with a 1 to a .000 value (i.e., for a value of .000, you would write it as .001).

Paired Samples Test								
Disability Group Membership	Paired Differences					t	df	Sig. (2-tailed)
				95% Confidence Interval of the Difference				
	Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
continued on next page								

²²<http://cnx.org/content/m37328/latest/figure5.8.1.PNG/image>

Students with Learning Disabilities	Pair 1	Verbal IQ (Wechsler Verbal Intelligence 3) - Performance 1 (Picture Completion)	76.192	10.141	.464	75.281	77.104	164.266	477	.000
-------------------------------------	--------	---	--------	--------	------	--------	--------	---------	-----	------

Table 6.2: Paired Samples Test

1. Numerical sentence is written as:

Numerical Sentence = $t(df)_{sp=sp} t_{sp} p_{sp < .001}$ (or Bonferroni-adjusted alpha).

- df is located in Paired Samples Box

- t is located in Paired Samples Box

2. The outcome of the paired samples t -test, $t(477) = 164.27$ $p < .001$, was statistically significant.

6.7 Step Five:**Check for Effect Size**

* Use the web-based calculator for effect size using the following websites:

Effect Size Calculators for Basic and Multivariate Statistical Procedures²³

Cohen's d (1988)

d of 0.20 = small effect size (range 0.20 to 0.49)

d of 0.50 = moderate effect size (range 0.50 to 0.79)

d of 0.80 = large effect size (range 0.80 and above)

Note. Cohen's d can be greater than 1.00. Therefore, a 0 should be placed in front of the decimal when the value is lower than 1.00.

²³<http://www.uccs.edu/~faculty/lbecker/>

Effect Size Calculators

Calculate Cohen's d and the effect-size correlation, r_{YA} , using --

- [means and standard deviations](#)
- [independent groups \$t\$ test values and \$df\$](#)

For a discussion of these effect size measures see [Effect Size Lecture Notes](#)

Calculate d and r using means and standard deviations

Calculate the value of Cohen's d and the effect-size correlation, r_{YA} , using the means and standard deviations of two groups (treatment and control).

Cohen's $d = M_1 - M_2 / \sigma_{\text{pooled}}$
 where $\sigma_{\text{pooled}} = \sqrt{[(\sigma_1^2 + \sigma_2^2) / 2]}$

$r_{YA} = d / \sqrt{d^2 + 4}$

Note: d and r_{YA} are positive if the mean difference is in the predicted direction.

Group 1	Group 2
M_1	M_2
SD_1	SD_2
Compute Reset	
Cohen's d	effect-size r

6.8 Step Six:

Narrative and Interpretation

1. type of t -test conducted and assumptions met
2. t value
3. degrees of freedom
4. p value

6.9 Writing Up Your Statistics

So, how do you "write up" your Research Questions and your Results? Schuler W. Huck (2000) in his seminal book entitled, *Reading Statistics and Research*, points to the importance of your audience understanding and making sense of your research in written form. Huck further states:

6.9.1

This book is designed to help people *decipher* what researchers are trying to communicate in the written or oral summaries of their investigations. Here, the goal is simply to distill meaning from the words, symbols, tables, and figures included in the research report. To be competent in this arena, one must not only be able to decipher what's presented but also to "fill in the holes"; this is the case because researchers typically assume that those receiving the research report are familiar with unmentioned details of the research process and statistical treatment of data.

Researchers and Professors John Slate and Ana Rojas-LeBouef understand this critical issue, so often neglected or not addressed by other authors and researchers. They point to the importance of doctoral students "writing up their statistics" in a way that others can understand your reporting and as importantly, interpret the meaning of your significant findings and implications for the preparation and practice of educational leadership. Slate and LeBouef provide you with a model for "writing up your parametric dependent sample t -test statistics."

Click here to view: *Writing Up Your Parametric Dependent Samples t-test Statistics* ²⁴

6.10 References

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum

Hyperstats Online Statistics Textbook. (n.d.) Retrieved from <http://davidmlane.com/hyperstat/> ²⁵

Kurtosis. (n.d.). Definition. Retrieved from http://www.statistics.com/index.php?page=glossary&term_id=326 ²⁶

Kurtosis. (n.d.). *Definition of normality*. Retrieved from <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb> ²⁷

Onwuegbuzie, A. J., & Daniel, L. G. (2002). Uses and misuses of the correlation coefficient. *Research in the Schools*, 9(1), 73-90.

Skewness. (n.d.) Retrieved from http://www.statistics.com/index.php?page=glossary&term_id=356 ²⁸

Skewness. (n.d.). *Definition of normality*. Retrieved from <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb> ²⁹

StatSoft, Inc. (2011). *Electronic statistics textbook*. Tulsa, OK: StatSoft. WEB: <http://www.statsoft.com/textbook/> ³⁰

²⁴<http://cnx.org/content/col11299/latest/>

²⁵<http://davidmlane.com/hyperstat/>

²⁶http://www.statistics.com/index.php?page=glossary&term_id=326

²⁷<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

²⁸http://www.statistics.com/index.php?page=glossary&term_id=356

²⁹<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

³⁰<http://www.statsoft.com/textbook/>

Chapter 7

Conducting a Nonparametric Independent Samples t-test



NOTE: This Chapter has been peer-reviewed, accepted, and endorsed by the National Council of Professors of Educational Administration (NCPEA) as a significant contribution to the scholarship and practice of education administration. Formatted and edited in Connexions by Theodore Creighton and Brad Bizzell, Virginia Tech, Janet Tareilo, Stephen F. Austin State University, and Thomas Kersten, Roosevelt University.

7.1

This chapter is part of a larger Collection (Book) and is available at: *Calculating Basic Statistical Procedures in SPSS: A Self-Help and Practical Guide to Preparing Theses, Dissertations, and Manuscripts* ²

NOTE: Slate and LeBouef have written a "companion book" which is available at: *Preparing and Presenting Your Statistical Findings: Model Write Ups* ³

Authors Information

John R. Slate is a Professor at Sam Houston State University where he teaches Basic and Advanced Statistics courses, as well as professional writing, to doctoral students in Educational Leadership and Counseling. His research interests lie in the use of educational databases, both state and national, to reform school practices. To date, he has chaired and/or served over 100 doctoral student dissertation committees. Recently, Dr. Slate created a website Writing and Statistical Help⁴ to assist students and faculty with both statistical assistance and in editing/writing their dissertations/theses and manuscripts.

¹This content is available online at <<http://cnx.org/content/m37330/1.6/>>.

²<http://my.qoop.com/store/NCPEA-Publications-1781472103076212/>

³<http://my.qoop.com/store/NCPEA-Publications-1781472103076212/>

⁴<http://www.writingandstatisticalhelp.com>

Ana Rojas-LeBouef is a Literacy Specialist at the Reading Center at Sam Houston State University where she teaches developmental reading courses. She recently completed her doctoral degree in Reading, where she conducted a 16-year analysis of Texas statewide data regarding the achievement gap. Her research interests lie in examining the inequities in achievement among ethnic groups. Dr. Rojas-LeBouef also assists students and faculty in their writing and statistical needs on the website Writing and Statistical Help.⁵

Editors Information

Theodore B. Creighton, is a Professor at Virginia Tech and the Publications Director for NCPEA Publications⁶, the Founding Editor of Education Leadership Review,⁷ and the Senior Editor of the NCPEA Connexions Project.

Brad E. Bizzell, is a recent graduate of the Virginia Tech Doctoral Program in Educational Leadership and Policy Studies, and is a School Improvement Coordinator for the Virginia Tech Training and Technical Assistance Center. In addition, Dr. Bizzell serves as an Assistant Editor of the NCPEA Connexions Project in charge of technical formatting and design.

Janet Tareilo, is a Professor at Stephen F. Austin State University and serves as the Assistant Director of NCPEA Publications. Dr. Tareilo also serves as an Assistant Editor of the NCPEA Connexions Project and as a editor and reviewer for several national and international journals in educational leadership.

Thomas Kersten is a Professor at Roosevelt University in Chicago. Dr. Kersten is widely published and an experienced editor and is the author of Taking the Mystery Out of Illinois School Finance⁸, a Connexions Print on Demand publication. He is also serving as Editor in Residence for this book by Slate and LeBouef.

7.2 Conducting a Nonparametric Independent Samples *t*-test

In this set of steps, readers will calculate either a parametric or a nonparametric statistical analysis, depending on whether the data for the dependent variable reflect a normal distribution. A parametric statistical procedure requires that its data be reflective of a normal curve whereas no such assumption is made in the use of a nonparametric procedure. Of the two types of statistical analyses, the parametric procedure is the more powerful one in ascertaining whether or not a statistically significant difference, in this case, exists. As such, parametric procedures are preferred over nonparametric procedures. When data are not normally distributed, however, parametric analyses may provide misleading and inaccurate results. According, nonparametric analyses should be used in cases where data are not reflective of a normal curve. In this set of steps, readers are provided with information on how to make the determination of normally or nonnormally distributed data. For detailed information regarding the assumptions underlying parametric and nonparametric procedures, readers are referred to the Hyperstats Online Statistics Textbook at <http://davidmlane.com/hyperstat/> or to the Electronic Statistics Textbook (2011) at <http://www.statsoft.com/textbook/>

For this nonparametric independent samples *t*-test to be appropriately used, at least half of the standardized skewness coefficients and the standardized kurtosis coefficients must be outside the normal range (± 3 , Onwuegbuzie & Daniel, 2002). Research questions for which nonparametric independent samples *t*-tests are appropriate involve asking for differences in a dependent variable by group membership (i.e., only two groups are present for *t*-tests). The research question, “What is the difference between boys and girls in their science performance among middle school students?” could be answered through use of a nonparametric independent samples *t*-test.

⁵<http://www.writingandstatisticalhelp.com>

⁶<http://www.ncpeapublications.org>

⁷<http://ncpeapublications.org/about-elr.html>

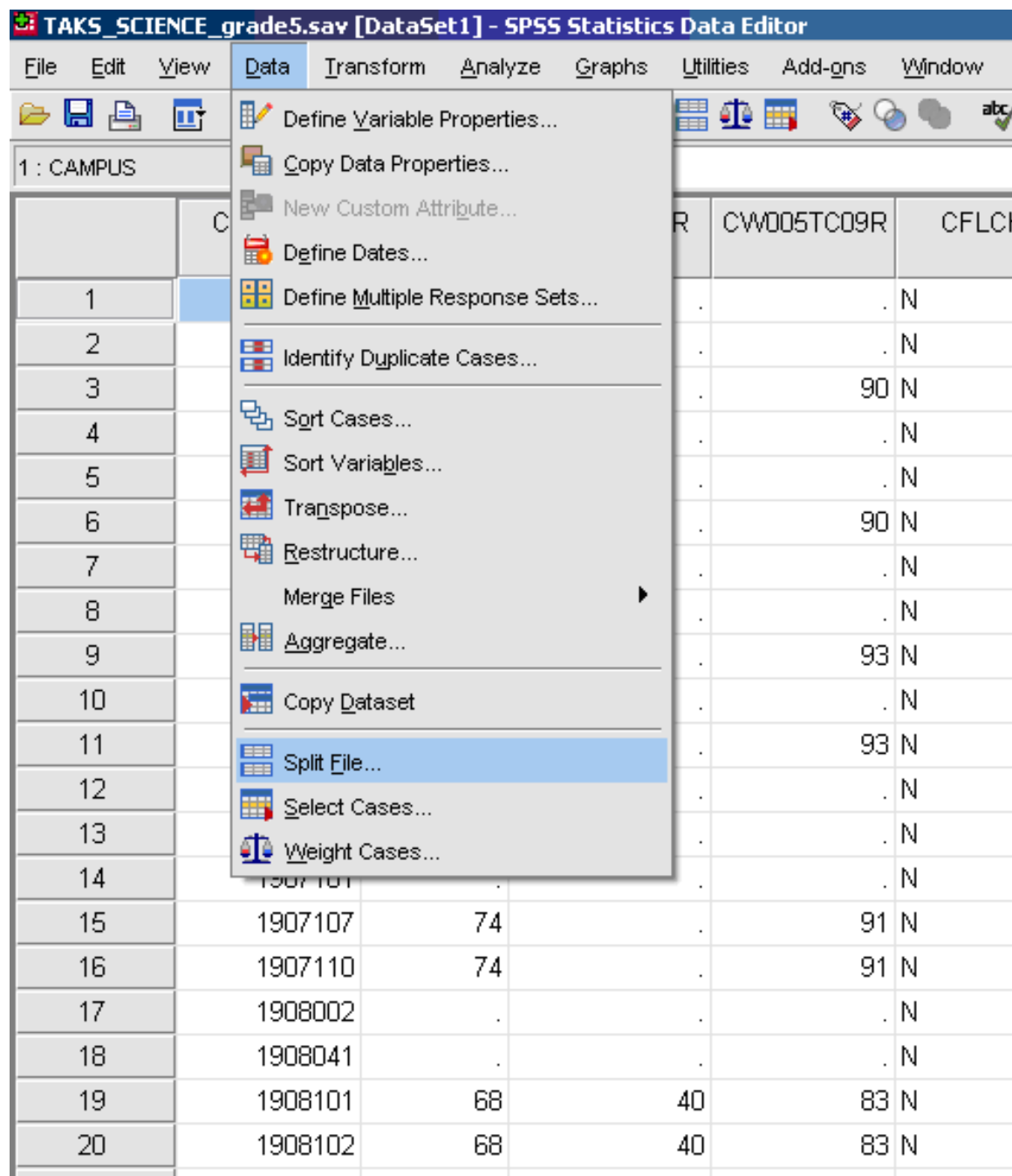
⁸<http://cnx.org/content/col10606/latest/>

7.3 Step One:

Calculate Frequencies on the Split Groups

✓ Data

* Split File



Your screen will show that all cases are going to be analyzed and a “do not create groups”. You will need to click the compare groups and move the independent variable over to the “Group Based on”.

⁹<http://cnx.org/content/m37330/latest/figure6.1.PNG/image>

SPSS Statistics Data Editor window showing the dataset 'TAKS_SCIENCE_grade5.sav [DataSet1]'. The data is organized by 'CAMPUS' (1902001 to 1902002). The 'Split File' dialog box is open, showing 'CAMPUS' as the grouping variable. The dialog has three radio buttons: 'Analyze all cases, do not create groups' (selected), 'Compare groups', and 'Organize output by groups'. Below these are options for 'Sort the file by grouping variables' and 'File is already sorted'. The status bar at the bottom of the dialog says 'Current Status: Analysis by groups is off.'

After you do this, your screen should resemble the following:

¹⁰<http://cnx.org/content/m37330/latest/figure6.2.PNG/image>

SPSS Statistics Data Editor window showing the dataset 'TAKS_SCIENCE_grade5.sav [DataSet1]'. The data is organized by 'CAMPUS' (1902001, 1902041, 1902103, 1903001, 1903041, 1903101, 1904001, 1904041, 1904102, 1906002, 1906102, 1907001, 1907041, 1907101, 1907107, 1907110, 1908002, 1908041, 1908101, 1908102, 1909001, 1909101, 2901001). The 'Split File' dialog box is open, showing 'CAMPUS' as the grouping variable. The 'Compare groups' option is selected. The 'Groups Based on:' list shows 'CH005TC09R' selected. The 'Current Status' at the bottom indicates 'Analysis by groups is off.'

11

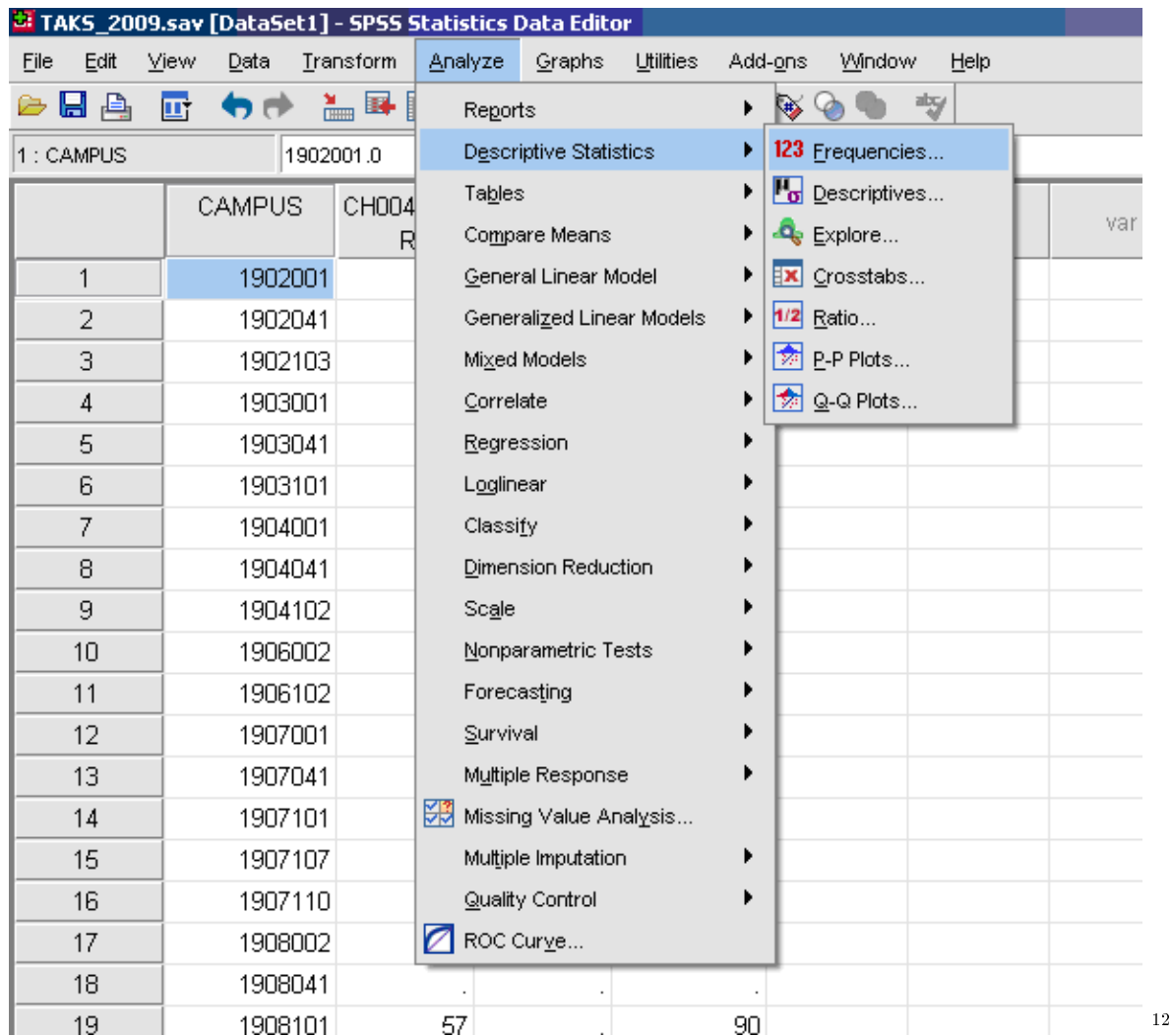
Then click OK

✓ Analyze

* Descriptive Statistics

* Frequencies

¹¹<http://cnx.org/content/m37330/latest/figure6.3.PNG/image>



✓ Move over the dependent (outcome) variable

¹²<http://cnx.org/content/m37330/latest/figure6.4.PNG/image>

1 : CAMPUS 1902001.0

	CAMPUS	CH005TC09R	CL005TC09R	CW005TC09R	CFLCHART	var	var	var
1	1902001	.	.	.	N			
2	1902041	.	.	.	N			
3	1902103	.	.	90	N			
4	1903001	.	.	.	N			
5	1903041	.	.	.	N			
6	1903101	.	.	90	N			
7	1904001	.	.	.	N			
8	1904041	.	.	.	N			
9	1904102	.	.	93	N			
10	1906002	.	.	.	N			
11	1906102	.	.	.				
12	1907001	.	.	.				
13	1907041	.	.	.				
14	1907101	.	.	.				
15	1907107	74	.	.				
16	1907110	74	.	.				
17	1908002	.	.	.				
18	1908041	.	.	.				
19	1908101	68	40	.				
20	1908102	68	40	.				
21	1909001	.	.	.				
22	1909101	.	.	.				
23	2901001	.	.	.	N			
24	2901002	.	.	.	N			

Frequencies

Variable(s): CH005TC09R

☐ Display frequency tables

OK Paste Reset Cancel Help

13

✓ Statistics

* Mean

* Standard Deviation

* Skewness [Note. Skewness refers to the extent to which the data are normally distributed around the mean. Skewed data involve having either mostly high scores with a few low ones or having mostly low scores with a few high ones.] Readers are referred to the following sources for a more detailed definition of skewness:

http://www.statistics.com/index.php?page=glossary&term_id=356¹⁴

<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>¹⁵

To standardize the skewness value so that its value can be constant across datasets and across studies, the following calculation must be made: Take the skewness value from the SPSS output and divide it by the Std. error of skewness. If the resulting calculation is within -3 to +3, then the skewness of the dataset is within the range of normality (Onwuegbuzie & Daniel, 2002). If the resulting calculation is outside of this +/-3 range, the dataset is not normally distributed.

* Kurtosis [Note. Kurtosis also refers to the extent to which the data are normally distributed around the mean. This time, the data are piled up higher than normal around the mean or piled up higher than normal at the ends of the distribution.] Readers are referred to the following sources for a more detailed definition of kurtosis:

To standardize the kurtosis value so that its value can be constant across datasets and across studies, the following calculation must be made: Take the kurtosis value from the SPSS output and divide it by the Std. error of kurtosis. If the resulting calculation is within -3 to +3, then the kurtosis of the dataset is within the range of normality (Onwuegbuzie & Daniel, 2002). If the resulting calculation is outside

¹³<http://cnx.org/content/m37330/latest/figure6.5.PNG/image>

¹⁴http://www.statistics.com/index.php?page=glossary&term_id=356

¹⁵<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

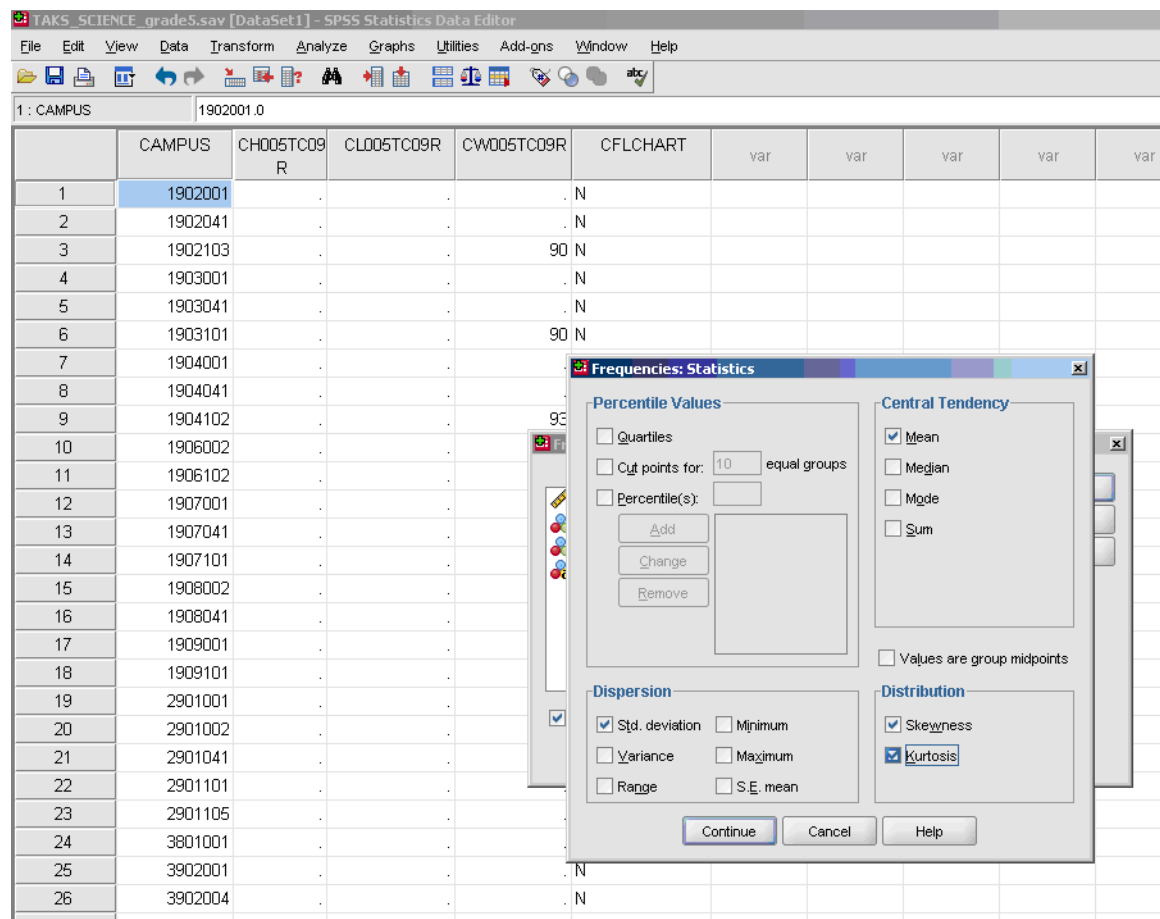
of this ± 3 range, the dataset is not normally distributed.

http://www.statistics.com/index.php?page=glossary&term_id=326¹⁶

<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>¹⁷

* Continue

* OK



18

✓ Charts (these are calculated only if you wish to have visual depictions of skewness and of kurtosis-they are not required)

* Histogram~ with normal curve (not required, optional)

✓ Continue

✓ OK

¹⁶http://www.statistics.com/index.php?page=glossary&term_id=326

¹⁷<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

¹⁸<http://cnx.org/content/m37330/latest/figure6.6.PNG/image>

The screenshot shows the SPSS Statistics Data Editor window with a data table. The table has columns: CAMPUS, CH004TR09R, CL004TR09R, CW004TR09R, and several 'var' columns. The data rows are numbered 1 to 25. Overlaid on the table are two dialog boxes. The 'Frequencies' dialog box has 'CAMPUS' selected in the variable list and 'Display frequencies' checked. The 'Frequencies: Charts' dialog box has 'None' selected under 'Chart Type' and 'Frequencies' selected under 'Chart Values'.

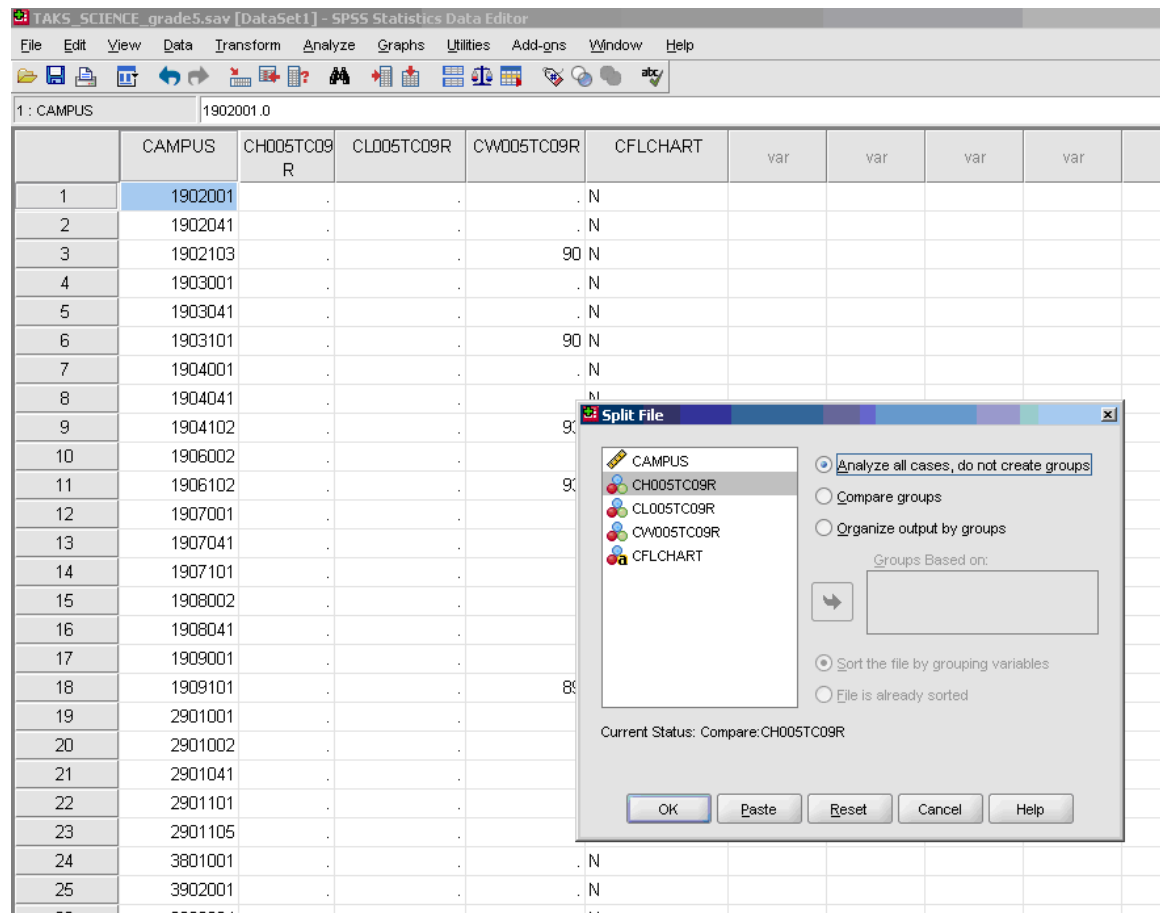
	CAMPUS	CH004TR09R	CL004TR09R	CW004TR09R	var	var	var	var	var	var	var
1	1902001	.	.	.							
2	1902041	.	.	.							
3	1902103	.	.	97							
4	1903001	.	.	.							
5	1903041	.	.	.							
6	1903101	89	.	97							
7	1904001	.	.	.							
8	1904041	.	.	.							
9	1904102	.	.	87							
10	1906002	.	.	.							
11	1906102	.	.	88							
12	1907001	.	.	.							
13	1907041	.	.	.							
14	1907101	.	.	.							
15	1907107	82	33	91							
16	1907110	82	33	91							
17	1908002	.	.	.							
18	1908041	.	.	.							
19	1908101	57	.	90							
20	1908102	57	.	90							
21	1909001	.	.	.							
22	1909101	.	.	70							
23	2901001	.	.	.							
24	2901002	.	.	.							
25	2901041	.	.	.							

19

Note: Before you continue to another application you must complete the following:

- ✓ Data
- ✓ Split Files
- ✓ Analyze all cases, do not create groups
- ✓ OK

¹⁹<http://cnx.org/content/m37330/latest/figure6.7.PNG/image>



20

7.4 Step Two:

Check for Skewness and Kurtosis values falling within/without the parameters of normality (-3 to +3). Note that each variable has its own skewness value and its own kurtosis value. Thus, a total of three standardized skewness coefficients and three standardized kurtosis coefficients can be calculated from information in the table below.

		CH005TC09R	CL005TC09R	CW005TC09R
N	Valid	3125	1805	1877
	Missing	5197	6517	6445

continued on next page

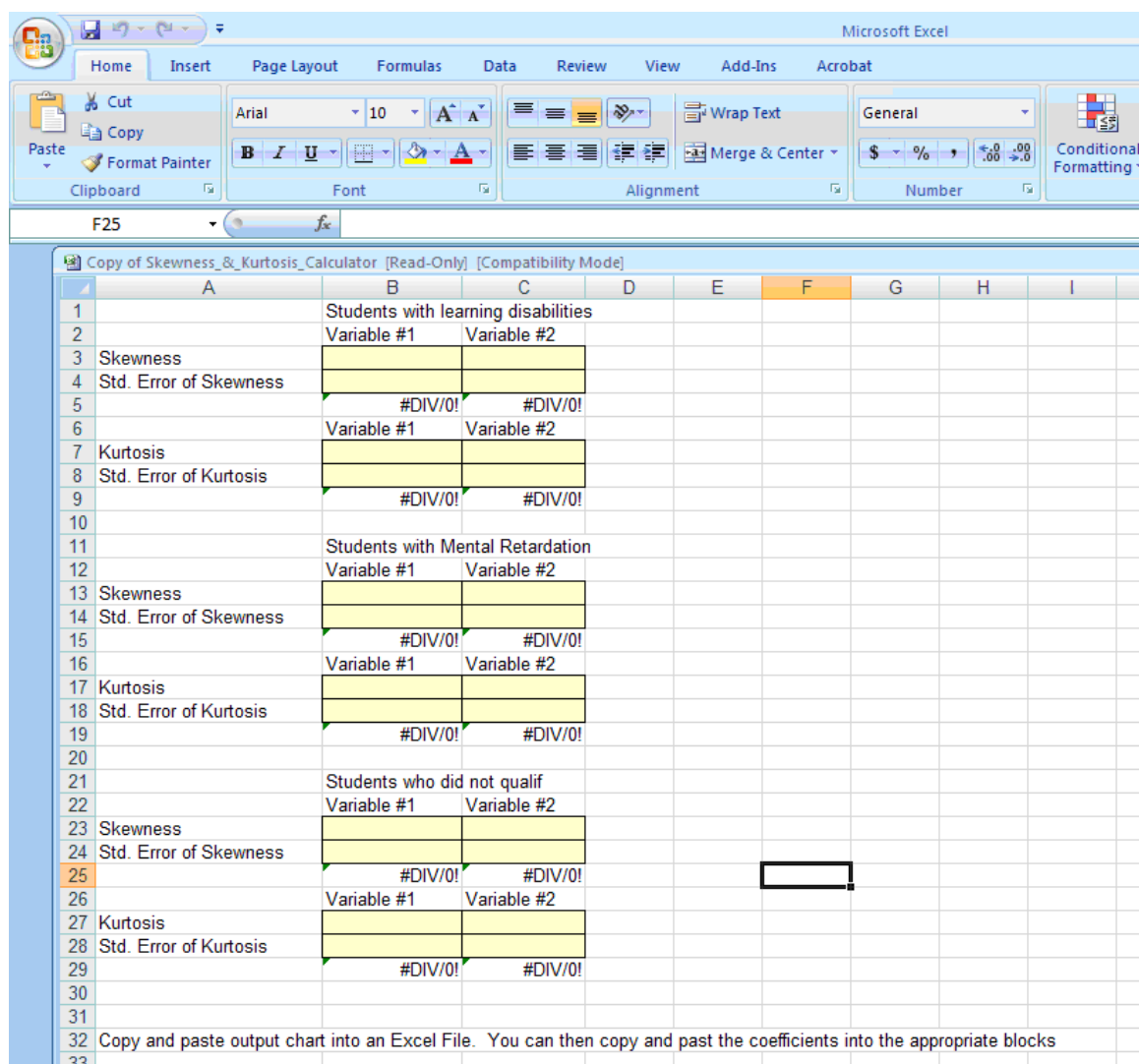
²⁰<http://cnx.org/content/m37330/latest/figure6.8.PNG/image>

Skewness	-1.129	-.479	-2.197
Std. Error of Skewness	.044	.058	.056
Kurtosis	1.818	-.412	6.991
Std. Error of Kurtosis	.088	.115	.113

Table 7.1: Skewness and Kurtosis Coefficients

Standard Coefficients Calculator

Copy variable #1 and #2 into the skewness and kurtosis calculator



21

7.5 Step Three

Calculate Nonparametric Independent Samples *t*-test on Data

²¹<http://cnx.org/content/m37330/latest/figure6.9.PNG/image>

- ✓ Analyze
- ✓ Nonparametric Tests
- ✓ 2 Independent Samples
- ✓ Test Variable would be your Dependent Variable (e.g., test scores)
- ✓ Grouping Variable would be your dichotomous Independent Variable

The screenshot shows the SPSS Statistics Data Editor window with the dataset 'Jamie complete spss dataset 07 year.sav [DataSet2]'. The data is organized into rows and columns. The first column is labeled '1 : CAMPUS' and contains values like 233901105, 233901106, etc. The second column is 'CPSTBLFP' with values like 0.04, 0.04, 0.08, etc. The third column is 'CA003TM07R' with values like 70, 80, 92, etc. The fourth column is 'CA003TR07R' with values like 81, 88, ., etc. The fifth column is 'AfAmTeachersGrouped' with values like ., ., ., etc. A dialog box titled 'Two-Independent-Samples Tests' is open in the foreground. It has a 'Test Variable List' on the right and a 'Grouping Variable' on the left. The 'Grouping Variable' is set to 'CAMPUS'. The 'Test Type' section has 'Mann-Whitney U' checked. There are buttons for 'Options...', 'Define Groups...', 'OK', 'Paste', 'Reset', 'Cancel', and 'Help'.

- ✓ Define Groups
 - ✓ Group One is No. 1 and Group Two is No. 2 (or whatever numbers you used to identify each group)
- Note:** Click on view than value labels to find the code for each group.
- ✓ Continue
 - ✓ OK

²²<http://cnx.org/content/m37330/latest/figure6.10.1.PNG/image>

The screenshot shows the SPSS Statistics Data Editor window with a dataset named 'Jamie complete spss dataset 07 year.sav [DataSet2]'. The data is organized into columns: CAMPUS, CPSTBLFP, CA003TM07R, CA003TR07R, AfAmTeachersGrouped, var, and v. A dialog box titled 'Two-Independent-Samples Tests' is open, allowing the user to select variables for testing. In the 'Test Variable List', 'CAMPUS' is selected. In the 'Grouping Variable' section, 'AfAmTeachersGrouped(? ?)' is selected. The 'Test Type' section shows 'Mann-Whitney U' as the selected test, with other options like 'Kolmogorov-Smirnov Z', 'Moses extreme reactions', and 'Wald-Wolfowitz runs' available but unchecked. Buttons for 'OK', 'Paste', 'Reset', 'Cancel', and 'Help' are at the bottom of the dialog.

23

7.6 Step Four:

Check for Statistical Significance

Test Statistics^a

	Performance IQ(Wechsler Performance Intelligence 3)
Mann-Whitney U	6765.500
Wilcoxon W	44166.500
Z	-20.752
Asymp. Sig. (2-tailed)	.000

Table 7.2

a. Grouping Variable:Disability Group Membership

Numerical sentence is written as: $U = 6765.50$, $p < .001$

²³<http://cnx.org/content/m37330/latest/figure6.11.1.PNG/image>

7.7 Step Five:

Check for Effect Size

* Use the web-based calculator for effect size using the following websites:

Effect Size Calculators for Basic and Multivariate Statistical Procedures²⁴

Effect Size Calculators

Calculate Cohen's d and the effect-size correlation, r_{YA} , using --

- [means and standard deviations](#)
- [independent groups \$t\$ test values and \$df\$](#)

For a discussion of these effect size measures see [Effect Size Lecture Notes](#)

Calculate d and r using means and standard deviations

Calculate the value of Cohen's d and the effect-size correlation, r_{YA} , using the means and standard deviations of two groups (treatment and control).

Cohen's $d = M_1 - M_2 / \sigma_{\text{pooled}}$
 where $\sigma_{\text{pooled}} = \sqrt{[(\sigma_1^2 + \sigma_2^2) / 2]}$

$r_{YA} = d / \sqrt{d^2 + 4}$

Note: d and r_{YA} are positive if the mean difference is in the predicted direction.

Group 1	Group 2
M_1	M_2
SD_1	SD_2
Compute Reset	
Cohen's d	effect-size r

25

7.8 Write Up Your Statistics

So, how do you "write up" your Research Questions and your Results? Schuler W. Huck (2000) in his seminal book entitled, *Reading Statistics and Research*, points to the importance of your audience understanding and making sense of your research in written form. Huck further states:

7.8.1

This book is designed to help people *decipher* what researchers are trying to communicate in the written or oral summaries of their investigations. Here, the goal is simply to distill meaning from the words, symbols, tables, and figures included in the research report. To be competent in this arena, one must not only be able to decipher what's presented but also to "fill in the holes"; this is the case because researchers typically assume that those receiving the research report are familiar with unmentioned details of the research process and statistical treatment of data.

Researchers and Professors John Slate and Ana Rojas-LeBouef understand this critical issue, so often neglected or not addressed by other authors and researchers. They point to the importance of doctoral students "writing up their statistics" in a way that others can understand your reporting and as importantly, interpret the meaning of your significant findings and implications for the preparation and practice of educational leadership. Slate and LeBouef provide you with a model for "writing up your nonparametric independent samples t -test statistics."

Click here to view: *Writing Up Your Nonparametric Independent Samples t -test Statistics* ²⁶

²⁴<http://www.uccs.edu/~faculty/lbecker/>

²⁵<http://cnx.org/content/m37330/latest/figure6.12.png/image>

²⁶<http://cnx.org/content/col11299/latest/>

7.9 References

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum
- Hyperstats Online Statistics Textbook. (n.d.) Retrieved from <http://davidmlane.com/hyperstat/>²⁷
- Kurtosis. (n.d.). Definition. Retrieved from http://www.statistics.com/index.php?page=glossary&term_id=326²⁸
- Kurtosis. (n.d.). *Definition of normality*. Retrieved from <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>²⁹
- Onwuegbuzie, A. J., & Daniel, L. G. (2002). Uses and misuses of the correlation coefficient. *Research in the Schools*, 9(1), 73-90.
- Skewness. (n.d.) Retrieved from http://www.statistics.com/index.php?page=glossary&term_id=356³⁰
- Skewness. (n.d.). *Definition of normality*. Retrieved from <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>³¹
- StatSoft, Inc. (2011). *Electronic statistics textbook*. Tulsa, OK: StatSoft. WEB: <http://www.statsoft.com/textbook/>³²

²⁷<http://davidmlane.com/hyperstat/>

²⁸http://www.statistics.com/index.php?page=glossary&term_id=326

²⁹<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

³⁰http://www.statistics.com/index.php?page=glossary&term_id=356

³¹<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

³²<http://www.statsoft.com/textbook/>

Chapter 8

Conducting a Nonparametric Paired Samples t-test



NOTE: This chapter has been peer-reviewed, accepted, and endorsed by the National Council of Professors of Educational Administration (NCPEA) as a significant contribution to the scholarship and practice of education administration. Formatted and edited in Connexions by Theodore Creighton and Brad Bizzell, Virginia Tech, Janet Tareilo, Stephen F. Austin State University, and Thomas Kersten, Roosevelt University.

8.1

This chapter is part of a larger Collection (Book) and is available at: *Calculating Basic Statistical Procedures in SPSS: A Self-Help and Practical Guide to Preparing Theses, Dissertations, and Manuscripts* ²

NOTE: Slate and LeBouef have written a "companion book" which is available at: *Preparing and Presenting Your Statistical Findings: Model Write Ups* ³

Authors Information

John R. Slate is a Professor at Sam Houston State University where he teaches Basic and Advanced Statistics courses, as well as professional writing, to doctoral students in Educational Leadership and Counseling. His research interests lie in the use of educational databases, both state and national, to reform school practices. To date, he has chaired and/or served over 100 doctoral student dissertation committees. Recently, Dr. Slate created a website Writing and Statistical Help⁴ to assist students and faculty with both statistical assistance and in editing/writing their dissertations/theses and manuscripts.

¹This content is available online at <<http://cnx.org/content/m37333/1.5/>>.

²<http://my.qoop.com/store/NCPEA-Publications-1781472103076212/>

³<http://my.qoop.com/store/NCPEA-Publications-1781472103076212/>

⁴<http://www.writingandstatisticalhelp.com>

Ana Rojas-LeBouef is a Literacy Specialist at the Reading Center at Sam Houston State University where she teaches developmental reading courses. She recently completed her doctoral degree in Reading, where she conducted a 16-year analysis of Texas statewide data regarding the achievement gap. Her research interests lie in examining the inequities in achievement among ethnic groups. Dr. Rojas-LeBouef also assists students and faculty in their writing and statistical needs on the website Writing and Statistical Help.⁵

Editors Information

Theodore B. Creighton, is a Professor at Virginia Tech and the Publications Director for NCPEA Publications⁶, the Founding Editor of Education Leadership Review,⁷ and the Senior Editor of the NCPEA Connexions Project.

Brad E. Bizzell, is a recent graduate of the Virginia Tech Doctoral Program in Educational Leadership and Policy Studies, and is a School Improvement Coordinator for the Virginia Tech Training and Technical Assistance Center. In addition, Dr. Bizzell serves as an Assistant Editor of the NCPEA Connexions Project in charge of technical formatting and design.

Janet Tareilo, is a Professor at Stephen F. Austin State University and serves as the Assistant Director of NCPEA Publications. Dr. Tareilo also serves as an Assistant Editor of the NCPEA Connexions Project and as a editor and reviewer for several national and international journals in educational leadership.

Thomas Kersten is a Professor at Roosevelt University in Chicago. Dr. Kersten is widely published and an experienced editor and is the author of Taking the Mystery Out of Illinois School Finance⁸, a Connexions Print on Demand publication. He is also serving as Editor in Residence for this book by Slate and LeBouef.

8.2 Conducting a Nonparametric Paired Samples t -test

In this set of steps, readers will calculate either a parametric or a nonparametric statistical analysis, depending on whether the data for the dependent variable reflect a normal distribution. A parametric statistical procedure requires that its data be reflective of a normal curve whereas no such assumption is made in the use of a nonparametric procedure. Of the two types of statistical analyses, the parametric procedure is the more powerful one in ascertaining whether or not a statistically significant difference, in this case, exists. As such, parametric procedures are preferred over nonparametric procedures. When data are not normally distributed, however, parametric analyses may provide misleading and inaccurate results. According, nonparametric analyses should be used in cases where data are not reflective of a normal curve. In this set of steps, readers are provided with information on how to make the determination of normally or nonnormally distributed data. For detailed information regarding the assumptions underlying parametric and nonparametric procedures, readers are referred to the Hyperstats Online Statistics Textbook at <http://davidmlane.com/hyperstat/> or to the Electronic Statistics Textbook (2011) at <http://www.statsoft.com/textbook/>

For this nonparametric dependent samples t -test to be appropriately used, at least half of the standardized skewness coefficients and the standardized kurtosis coefficients must be outside the normal range (± 3 , Onwuegbuzie & Daniel, 2002). Research questions for which nonparametric dependent samples t -test are appropriate involve asking for differences in a dependent variable by group membership (i.e., only two groups are present for the t -test and, in this case, their scores are connected). The research question, “What is the effect of the new science program on student science performance among elementary school students?” could be answered through use of a nonparametric dependent dependent t -test.

⁵<http://www.writingandstatisticalhelp.com>

⁶<http://www.ncpeapublications.org>

⁷<http://ncpeapublications.org/about-elr.html>

⁸<http://cnx.org/content/col10606/latest/>

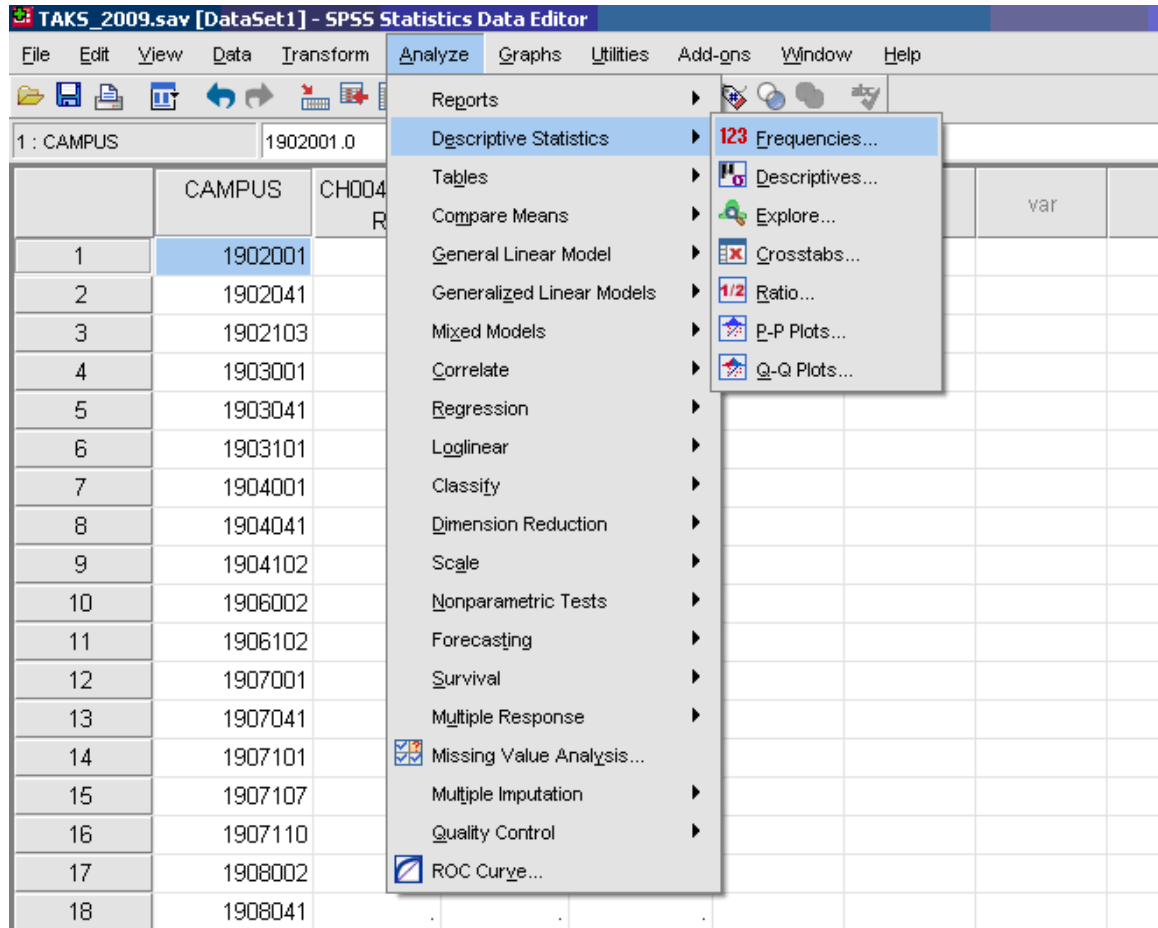
8.2.1 Step One:

Compute Measures of Normality for the Dependent Variable

✓ Analyze

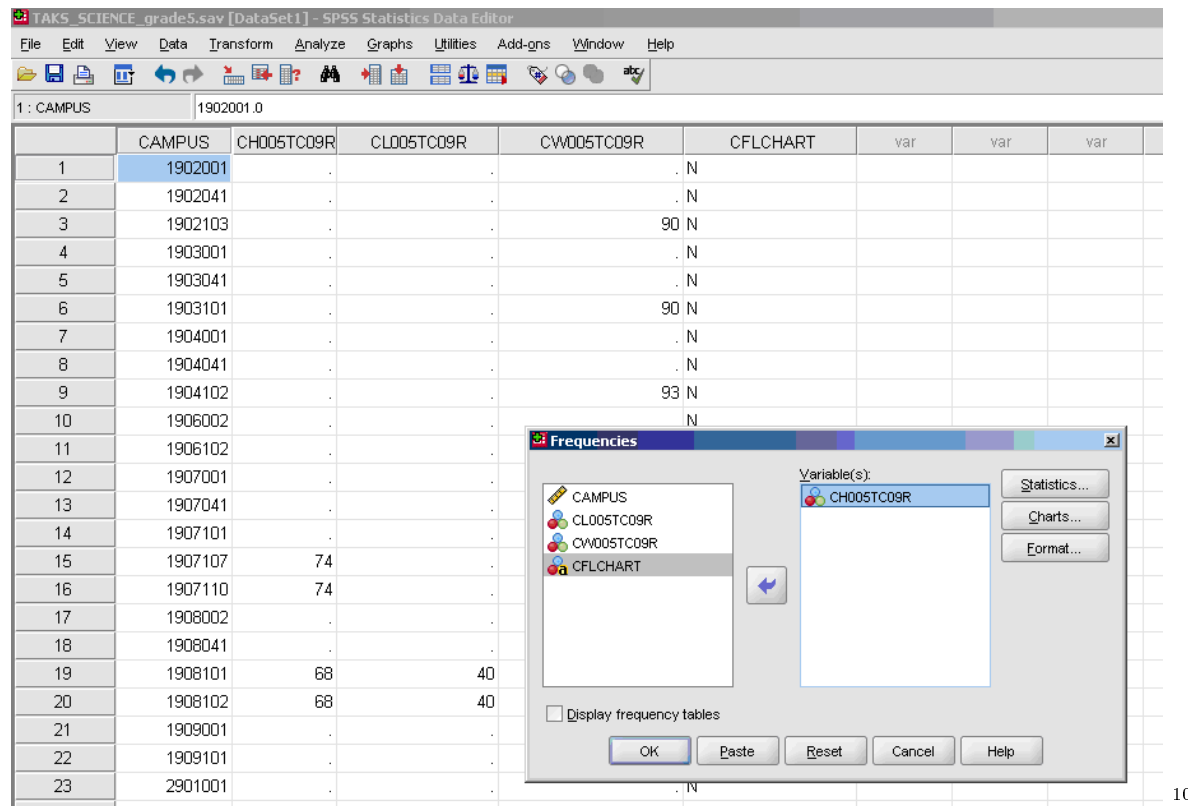
* Descriptive Statistics

* Frequencies



✓ Move over the dependent (outcome) variable

⁹<http://cnx.org/content/m37333/latest/figure7.1.PNG/image>



✓ Statistics

* Skewness [Note. Skewness refers to the extent to which the data are normally distributed around the mean. Skewed data involve having either mostly high scores with a few low ones or having mostly low scores with a few high ones.] Readers are referred to the following sources for a more detailed definition of skewness: http://www.statistics.com/index.php?page=glossary&term_id=356 and <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

To standardize the skewness value so that its value can be constant across datasets and across studies, the following calculation must be made: Take the skewness value from the SPSS output and divide it by the Std. error of skewness. If the resulting calculation is within -3 to +3, then the skewness of the dataset is within the range of normality (Onwuegbuzie & Daniel, 2002). If the resulting calculation is outside of this +/-3 range, the dataset is not normally distributed.

* Kurtosis [Note. Kurtosis also refers to the extent to which the data are normally distributed around the mean. This time, the data are piled up higher than normal around the mean or piled up higher than normal at the ends of the distribution.] Readers are referred to the following sources for a more detailed definition of kurtosis: http://www.statistics.com/index.php?page=glossary&term_id=326 and <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

To standardize the kurtosis value so that its value can be constant across datasets and across studies, the following calculation must be made: Take the kurtosis value from the SPSS output and divide it by the Std. error of kurtosis. If the resulting calculation is within -3 to +3, then the kurtosis of the dataset is within the range of normality (Onwuegbuzie & Daniel, 2002). If the resulting calculation is outside of this +/-3 range, the dataset is not normally distributed.

* Continue

* OK

¹⁰<http://cnx.org/content/m37333/latest/figure7.2.PNG/image>

The screenshot shows the SPSS Statistics Data Editor window with a dataset named 'TAKS_2009.sav'. The dataset has columns for 'CAMPUS', 'CH004TR09R', 'CL004TR09R', and 'CW004TR09R'. A 'Frequencies: Statistics' dialog box is open, showing options for Percentile Values, Central Tendency, Dispersion, and Distribution. The 'Distribution' section has 'Skewness' and 'Kurtosis' checked.

11

8.2.2 Step Two:

Check for Skewness and Kurtosis values falling within/without the parameters of normality (-3 to +3). Note that each variable below has its own skewness value and its own kurtosis value. Thus, a total of three standardized skewness coefficients and three standardized kurtosis coefficients can be calculated from information in the table below.

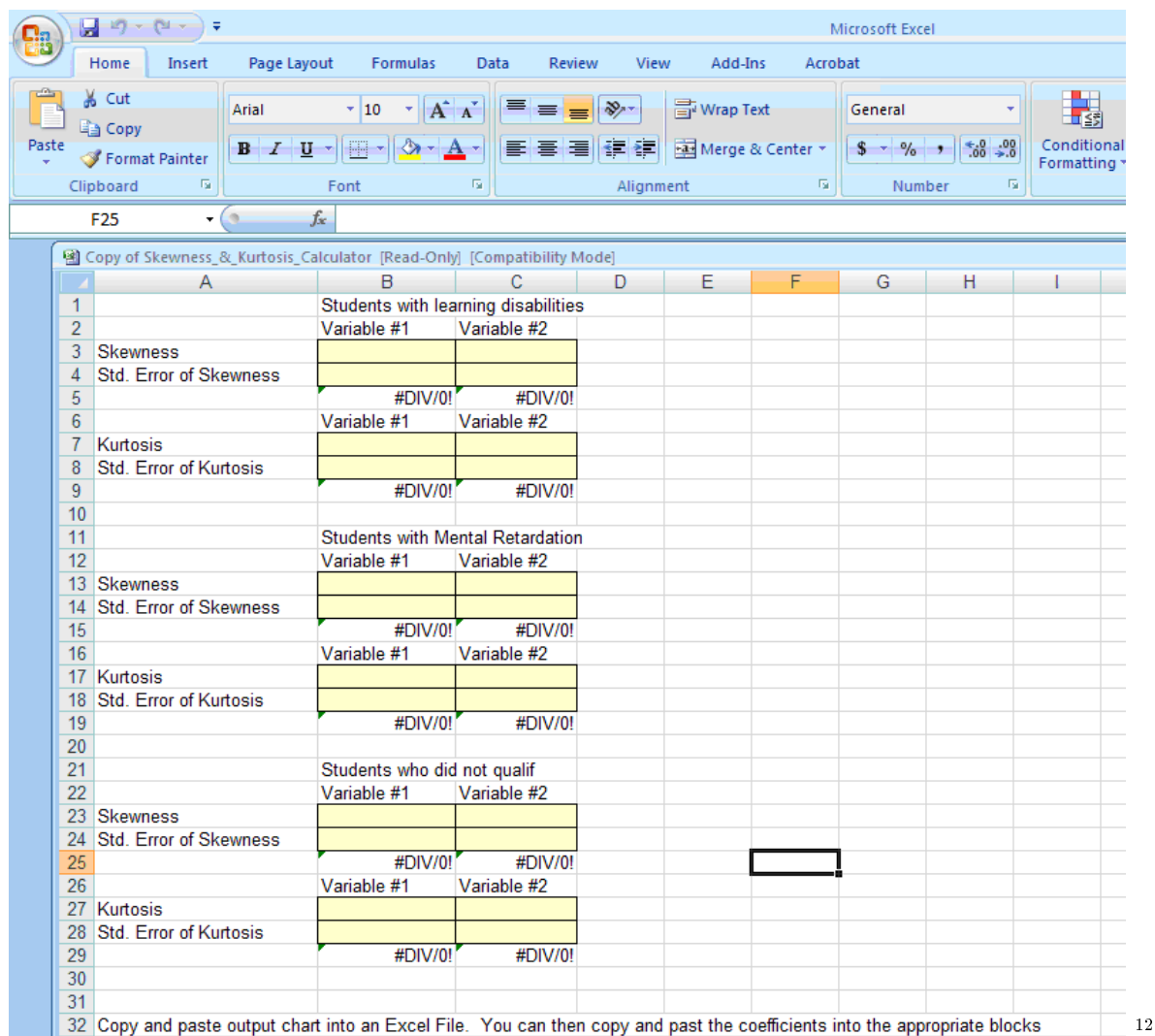
		CH005TC09R	CL005TC09R	CW005TC09R
N	Valid	3125	1805	1877
Missing	5197	6517	6445	
Skewness		-1.129	-.479	-2.197
Std. Error of Skewness		.044	.058	.056
Kurtosis		1.818	-.412	6.991
Std. Error of Kurtosis		.088	.115	.113

Table 8.1: Skewness and Kurtosis coefficients

¹¹<http://cnx.org/content/m37333/latest/figure7.3.PNG/image>

Standard Coefficients Calculator

Copy variable #1 and #2 into the skewness and kurtosis calculator



✓ Charts (these are calculated only if you wish to have visual depictions of skewness and of kurtosis-they are not required)

* Histogram~ with normal curve (not required, optional)

¹²<http://cnx.org/content/m37333/latest/figure7.4.PNG/image>

23:

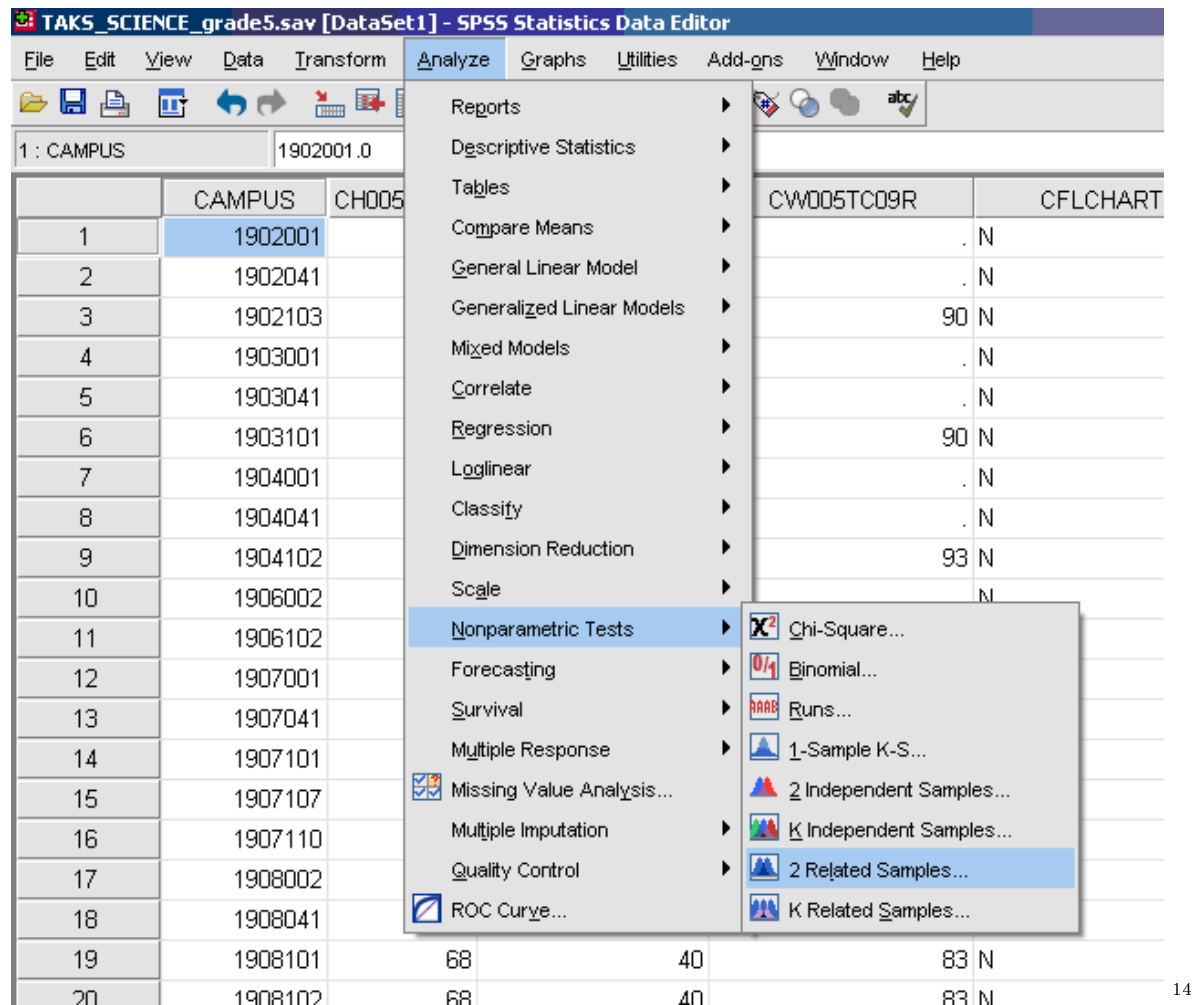
	CAMPUS	CH004TR09 R	CL004TR09 R	CW004TR09F	var	var	var	var	var	var	var
1	1902001	.	.	.							
2	1902041	.	.	.							
3	1902103	.	.	97							
4	1903001	.	.	.							
5	1903041	.	.	.							
6	1903101	89	.	97							
7	1904001	.	.	.							
8	1904041	.	.	.							
9	1904102	.	.	87							
10	1906002	.	.	.							
11	1906102	.	.	88							
12	1907001	.	.	.							
13	1907041	.	.	.							
14	1907101	.	.	.							
15	1907107	82	33	91							
16	1907110	82	33	91							
17	1908002	.	.	.							
18	1908041	.	.	.							
19	1908101	57	.	90							
20	1908102	57	.	90							
21	1909001	.	.	.							
22	1909101	.	.	70							
23	2901001	.	.	.							
24	2901002	.	.	.							

13

8.2.3 Step Three:

Calculate Nonparametric Paired Samples t -test on Data

¹³<http://cnx.org/content/m37333/latest/figure7.5.PNG/image>



- ✓ Click on one dependent variable
- * Arrow to send over to Test Pairs, Variable 1
- ✓ Click on second dependent variable
- * Arrow to send over to Test Pairs, Variable 2
- * OK

¹⁴<http://cnx.org/content/m37333/latest/figure7.6.PNG/image>

1 : CAMPUS 1902001.0

	CAMPUS	CH005TC09R	CL005TC09R	CW005TC09R	CFLCHART	var	var	var	var
1	1902001	.	.	.	N				
2	1902041	.	.	.	N				
3	1902103	.	.	90	N				
4	1903001	.	.	.	N				
5	1903041	.	.	.	N				
6	1903101	.	.	90	N				
7	1904001	.	.	.	N				
8	1904041	.	.	.	N				
9	1904102	.	.	.					
10	1906002	.	.	.					
11	1906102	.	.	.					
12	1907001	.	.	.					
13	1907041	.	.	.					
14	1907101	.	.	.					
15	1907107	74	.	.					
16	1907110	74	.	.					
17	1908002	.	.	.					
18	1908041	.	.	.					
19	1908101	68	.	.					
20	1908102	68	.	.					
21	1909001	.	.	.					
22	1909101	.	.	.					
23	2901001	.	.	.					
24	2901002	.	.	.					
25	2901041	.	.	.	N				
26	2901101	.	.	.	N				

Two-Related-Samples Tests

Test Pairs:

Pair	Variable1	Variable2
1	[CH005TC09R]	[CL005TC09R]
2		

Test Type

☒ Wilcoxon

☐ Sign

☐ McNemar

☐ Marginal Homogeneity

OK Paste Reset Cancel Help

15

8.2.4 Step Four:

Check for Statistical Significance

Test Statistics^b

	CL005TC09R
	CH005TC09R
Z	-34.829 ^a
Asymp. Sig. (2-tailed)	.000

Table 8.2

a. Based on positive ranks.

b. Wilcoxon Signed Rank test

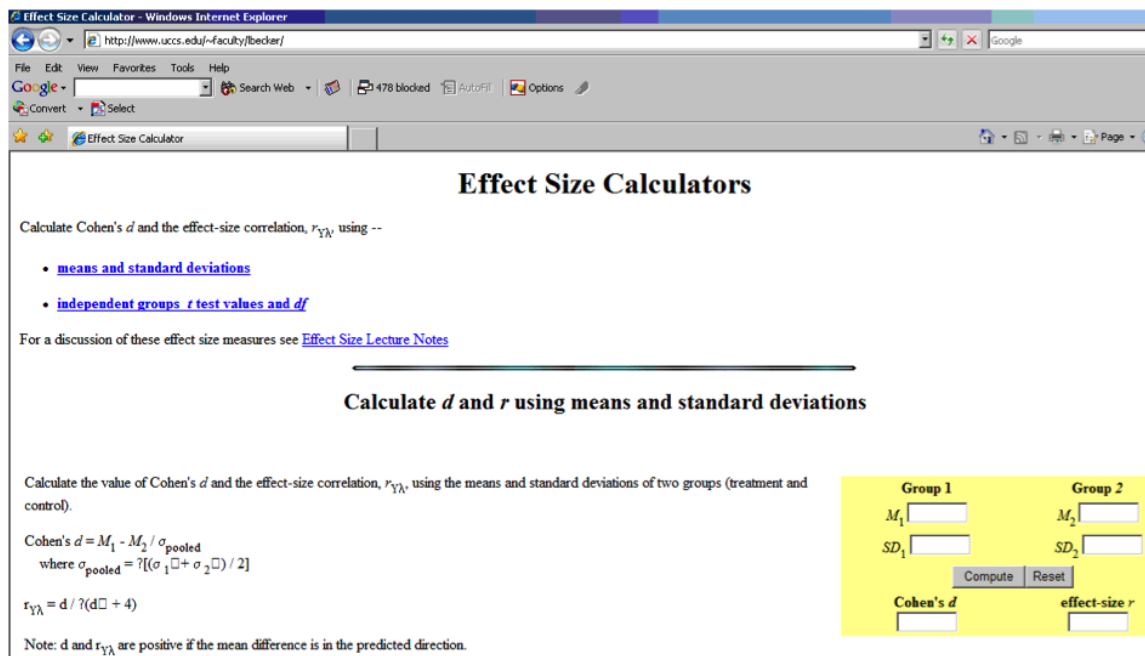
Numerical sentence is written as: $z = -34.83, p < .001$

¹⁵<http://cnx.org/content/m37333/latest/figure7.7.PNG/image>

8.2.5 Step Five

Check for Effect Size

* Use the web-based calculator for effect size using the following website:
Effect Size Calculators for Basic and Multivariate Statistical Procedures¹⁶



Effect Size Calculators

Calculate Cohen's d and the effect-size correlation, $r_{Y\lambda}$, using --

- [means and standard deviations](#)
- [independent groups \$t\$ test values and \$df\$](#)

For a discussion of these effect size measures see [Effect Size Lecture Notes](#)

Calculate d and r using means and standard deviations

Calculate the value of Cohen's d and the effect-size correlation, $r_{Y\lambda}$, using the means and standard deviations of two groups (treatment and control).

Cohen's $d = M_1 - M_2 / \sigma_{\text{pooled}}$
 where $\sigma_{\text{pooled}} = \sqrt{[(\sigma_1^2 + \sigma_2^2) / 2]}$

$r_{Y\lambda} = d / \sqrt{(d^2 + 4)}$

Note: d and $r_{Y\lambda}$ are positive if the mean difference is in the predicted direction.

Group 1	Group 2
M_1	M_2
SD_1	SD_2
Compute Reset	
Cohen's d	effect-size r

17

8.3 Writing Up Your Statistics

So, how do you "write up" your Research Questions and your Results? Schuler W. Huck (2000) in his seminal book entitled, *Reading Statistics and Research*, points to the importance of your audience understanding and making sense of your research in written form. Huck further states:

8.3.1

This book is designed to help people *decipher* what researchers are trying to communicate in the written or oral summaries of their investigations. Here, the goal is simply to distill meaning from the words, symbols, tables, and figures included in the research report. To be competent in this arena, one must not only be able to decipher what's presented but also to "fill in the holes"; this is the case because researchers typically assume that those receiving the research report are familiar with unmentioned details of the research process and statistical treatment of data.

Researchers and Professors John Slate and Ana Rojas-LeBouef understand this critical issue, so often neglected or not addressed by other authors and researchers. They point to the importance of doctoral students "writing up their statistics" in a way that others can understand your reporting and as importantly, interpret the meaning of your significant findings and implications for the preparation and practice of educational leadership. Slate and LeBouef provide you with a model for "writing up your nonparametric paired samples t-test."

Click here to view: *Writing Up Your Nonparametric Paired Samples t-test Statistics*¹⁸

¹⁶<http://www.uccs.edu/~faculty/lbecker/>

¹⁷<http://cnx.org/content/m37333/latest/figure7.8.png/image>

¹⁸<http://cnx.org/content/col11299/latest/>

8.4 References

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum
- Hyperstats Online Statistics Textbook. (n.d.) Retrieved from <http://davidmlane.com/hyperstat/>¹⁹
- Kurtosis. (n.d.). Definition. Retrieved from http://www.statistics.com/index.php?page=glossary&term_id=326²⁰
- Kurtosis. (n.d.). *Definition of normality*. Retrieved from <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>²¹
- Onwuegbuzie, A. J., & Daniel, L. G. (2002). Uses and misuses of the correlation coefficient. *Research in the Schools*, 9(1), 73-90.
- Skewness. (n.d.) Retrieved from http://www.statistics.com/index.php?page=glossary&term_id=356²²
- Skewness. (n.d.). *Definition of normality*. Retrieved from <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>²³
- StatSoft, Inc. (2011). *Electronic statistics textbook*. Tulsa, OK: StatSoft. WEB: <http://www.statsoft.com/textbook/>²⁴

¹⁹<http://davidmlane.com/hyperstat/>

²⁰http://www.statistics.com/index.php?page=glossary&term_id=326

²¹<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

²²http://www.statistics.com/index.php?page=glossary&term_id=356

²³<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

²⁴<http://www.statsoft.com/textbook/>

Chapter 9

Conducting a Parametric One-Way Analysis of Variance



NOTE: This chapter has been peer-reviewed, accepted, and endorsed by the National Council of Professors of Educational Administration (NCPEA) as a significant contribution to the scholarship and practice of education administration. Formatted and edited in Connexions by Theodore Creighton and Brad Bizzell, Virginia Tech, Janet Tareilo, Stephen F. Austin State University, and Thomas Kersten, Roosevelt University.

9.1

This chapter is part of a larger Collection (Book) and is available at: *Calculating Basic Statistical Procedures in SPSS: A Self-Help and Practical Guide to Preparing Theses, Dissertations, and Manuscripts*²

NOTE: Slate and LeBouef have written a "companion book" which is available at: *Preparing and Presenting Your Statistical Findings: Model Write Ups*³

Authors Information

John R. Slate is a Professor at Sam Houston State University where he teaches Basic and Advanced Statistics courses, as well as professional writing, to doctoral students in Educational Leadership and Counseling. His research interests lie in the use of educational databases, both state and national, to reform school practices. To date, he has chaired and/or served over 100 doctoral student dissertation committees. Recently, Dr. Slate created a website Writing and Statistical Help⁴ to assist students and faculty with both statistical assistance and in editing/writing their dissertations/theses and manuscripts.

¹This content is available online at <<http://cnx.org/content/m37334/1.4/>>.

²<http://my.qoop.com/store/NCPEA-Publications-1781472103076212/>

³<http://my.qoop.com/store/NCPEA-Publications-1781472103076212/>

⁴<http://www.writingandstatisticalhelp.com>

Ana Rojas-LeBouef is a Literacy Specialist at the Reading Center at Sam Houston State University where she teaches developmental reading courses. She recently completed her doctoral degree in Reading, where she conducted a 16-year analysis of Texas statewide data regarding the achievement gap. Her research interests lie in examining the inequities in achievement among ethnic groups. Dr. Rojas-LeBouef also assists students and faculty in their writing and statistical needs on the website Writing and Statistical Help.⁵

Editors Information

Theodore B. Creighton, is a Professor at Virginia Tech and the Publications Director for NCPEA Publications⁶, the Founding Editor of Education Leadership Review,⁷ and the Senior Editor of the NCPEA Connexions Project.

Brad E. Bizzell, is a recent graduate of the Virginia Tech Doctoral Program in Educational Leadership and Policy Studies, and is a School Improvement Coordinator for the Virginia Tech Training and Technical Assistance Center. In addition, Dr. Bizzell serves as an Assistant Editor of the NCPEA Connexions Project in charge of technical formatting and design.

Janet Tareilo, is a Professor at Stephen F. Austin State University and serves as the Assistant Director of NCPEA Publications. Dr. Tareilo also serves as an Assistant Editor of the NCPEA Connexions Project and as a editor and reviewer for several national and international journals in educational leadership.

Thomas Kersten is a Professor at Roosevelt University in Chicago. Dr. Kersten is widely published and an experienced editor and is the author of Taking the Mystery Out of Illinois School Finance⁸, a Connexions Print on Demand publication. He is also serving as Editor in Residence for this book by Slate and LeBouef.

9.2 Conducting a Parametric One-Way Analysis of Variance

In this set of steps, readers will calculate either a parametric or a nonparametric statistical analysis, depending on whether the data for the dependent variable reflect a normal distribution. A parametric statistical procedure requires that its data be reflective of a normal curve whereas no such assumption is made in the use of a nonparametric procedure. Of the two types of statistical analyses, the parametric procedure is the more powerful one in ascertaining whether or not a statistically significant difference, in this case, exists. As such, parametric procedures are preferred over nonparametric procedures. When data are not normally distributed, however, parametric analyses may provide misleading and inaccurate results. According, nonparametric analyses should be used in cases where data are not reflective of a normal curve. In this set of steps, readers are provided with information on how to make the determination of normally or nonnormally distributed data. For detailed information regarding the assumptions underlying parametric and nonparametric procedures, readers are referred to the Hyperstats Online Statistics Textbook at <http://davidmlane.com/hyperstat/> or to the Electronic Statistics Textbook (2011) at <http://www.statsoft.com/textbook/>

For this parametric analysis of variance procedure to be appropriately used, at least half of the standardized skewness coefficients and the standardized kurtosis coefficients must be within the normal range (± 3 , Onwuegbuzie & Daniel, 2002). Research questions for which parametric analysis of variance procedures are appropriate involve asking for differences in a dependent variable by group membership (i.e., more than two groups may be present). The research question, “What is the difference in science achievement among elementary school students as a function of ethnic membership?” could be answered through use of an analysis of variance procedure.

⁵<http://www.writingandstatisticalhelp.com>

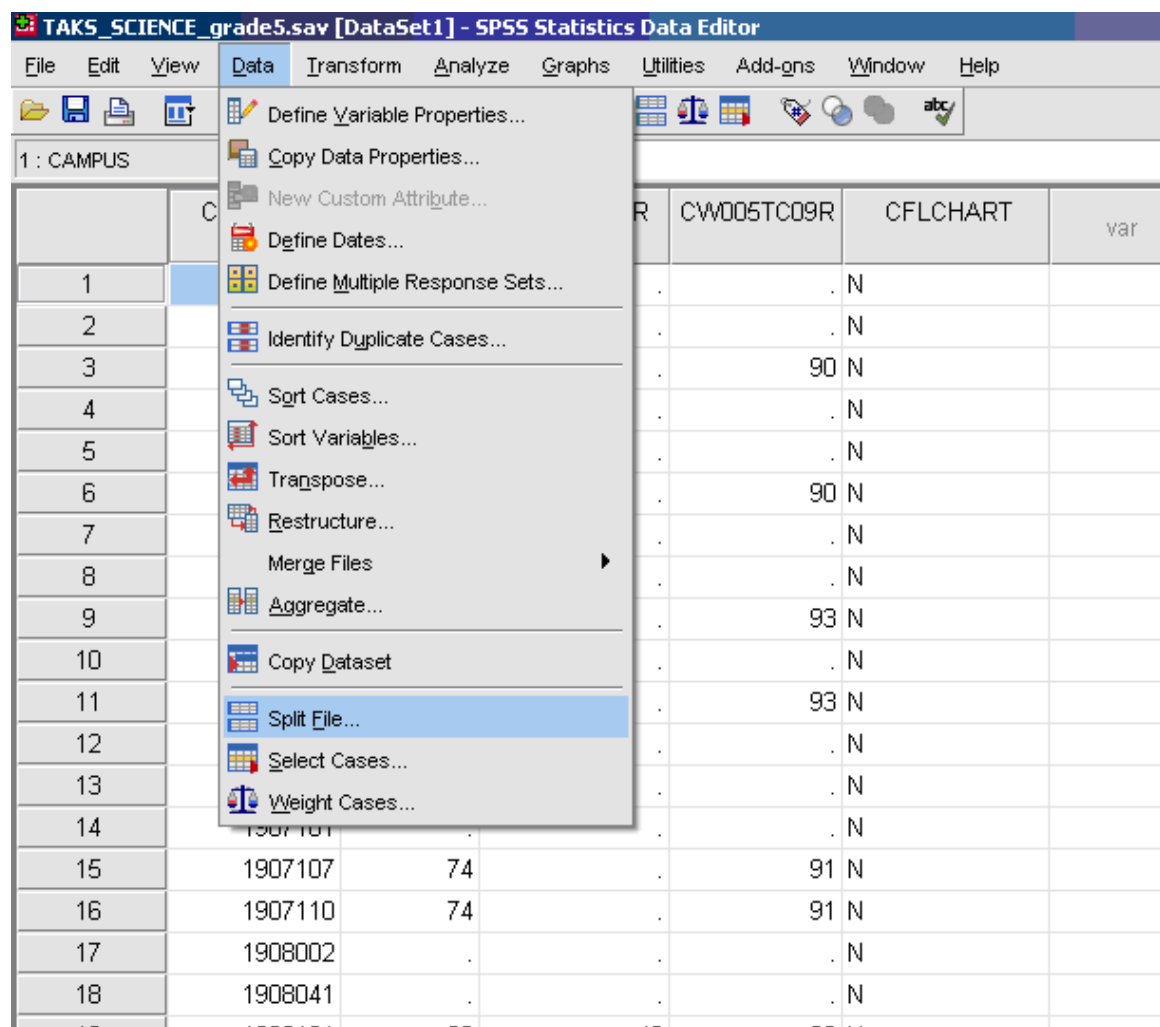
⁶<http://www.ncpeapublications.org>

⁷<http://ncpeapublications.org/about-elr.html>

⁸<http://cnx.org/content/col10606/latest/>

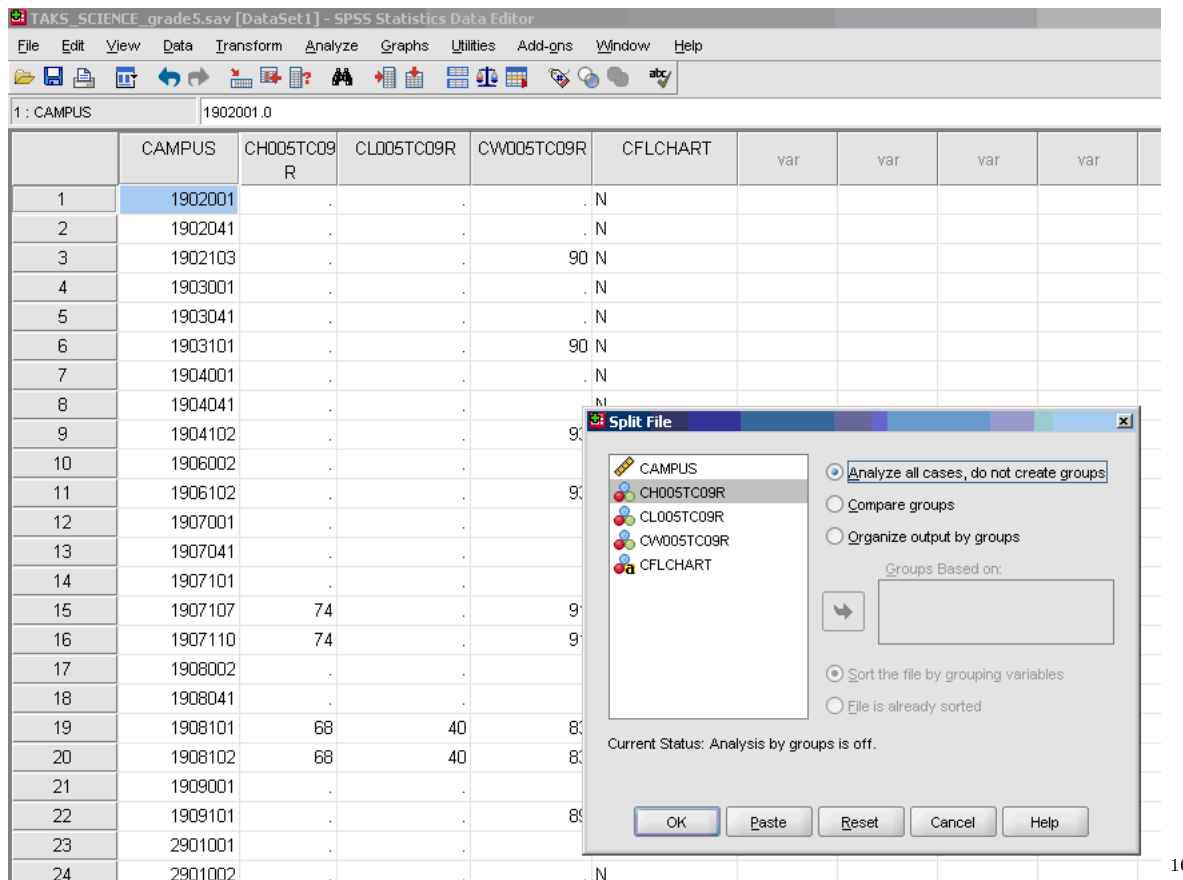
9.2.1 Step One:

Check for Skewness and Kurtosis values falling within/without the parameters of normality (-3 to +3)
 ✓ Split your file on the basis on your independent variable/fixed factor/grouping variable



After you do this, your screen should resemble the following:

⁹<http://cnx.org/content/m37334/latest/figure8.1.PNG/image>



Your screen will show that all cases are going to be analyzed and a “do not create groups”. You will need to click the compare groups and move the independent variable over to the “Group Based on”. For most ANOVA procedures, your independent or grouping variable will have more than two groups.

¹⁰<http://cnx.org/content/m37334/latest/figure8.2.1.PNG/image>

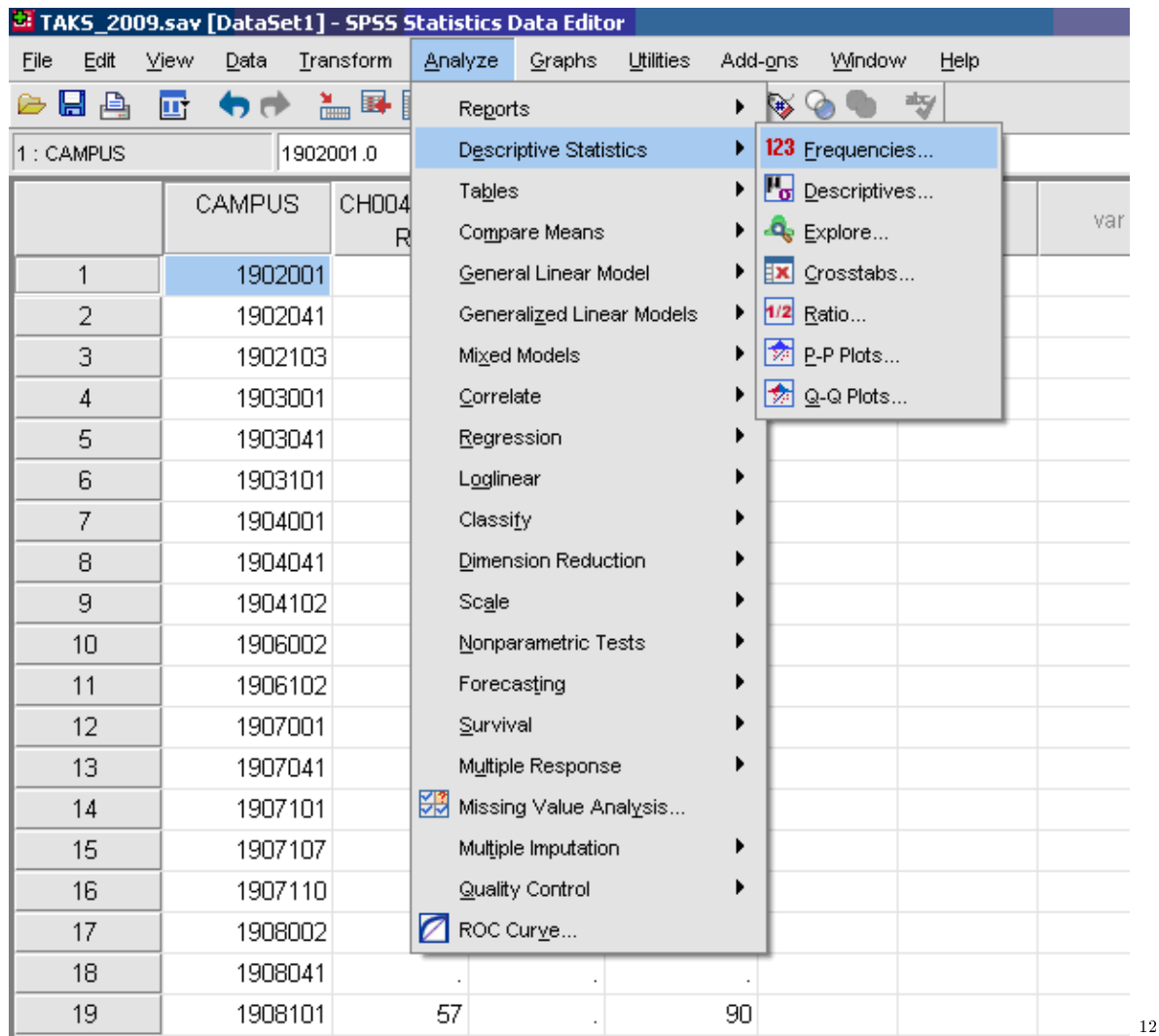
The screenshot shows the SPSS Statistics Data Editor window with a dataset named 'practice SPSS database revised.sav [DataSet1]'. The dataset has 24 rows and 10 columns: group, id, age, recodage, wifsiq, wiviq, wipiq, pc, inf, and c. The 'Split File' dialog box is open, showing the following options:

- ☐ Analyze all cases, do not create groups
- ☒ Compare groups
- ☐ Organize output by groups

Under 'Groups Based on:', the variable 'Disability Group Membership' is selected. The 'Sort the file by grouping variables' option is also selected. The 'File is already sorted' option is not selected. The 'Current Status: Analysis by groups is off.' message is displayed at the bottom of the dialog box.

- ✓ Analyze
- * Descriptive Statistics
- * Frequencies

¹¹ <http://cnx.org/content/m37334/latest/figure8.3.PNG/image>



✓ Move over the dependent (outcome) variable

¹²<http://cnx.org/content/m37334/latest/figure8.4.PNG/image>

TAKS_SCIENCE_grade5.sav [DataSet1] - SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

1 : CAMPUS 1902001.0

	CAMPUS	CH005TC09R	CL005TC09R	CW005TC09R	CFLCHART	var	var	var
1	1902001	.	.	.	N			
2	1902041	.	.	.	N			
3	1902103	.	.	90	N			
4	1903001	.	.	.	N			
5	1903041	.	.	.	N			
6	1903101	.	.	90	N			
7	1904001	.	.	.	N			
8	1904041	.	.	.	N			
9	1904102	.	.	93	N			
10	1906002	.	.	.	N			
11	1906102	.	.	.				
12	1907001	.	.	.				
13	1907041	.	.	.				
14	1907101	.	.	.				
15	1907107	74	.	.				
16	1907110	74	.	.				
17	1908002	.	.	.				
18	1908041	.	.	.				
19	1908101	68	40	.				
20	1908102	68	40	.				
21	1909001	.	.	.				
22	1909101	.	.	.				
23	2901001	.	.	.	N			
24	2901002	.	.	.	N			

Frequencies

Variable(s): CH005TC09R

Statistics... Charts... Format...

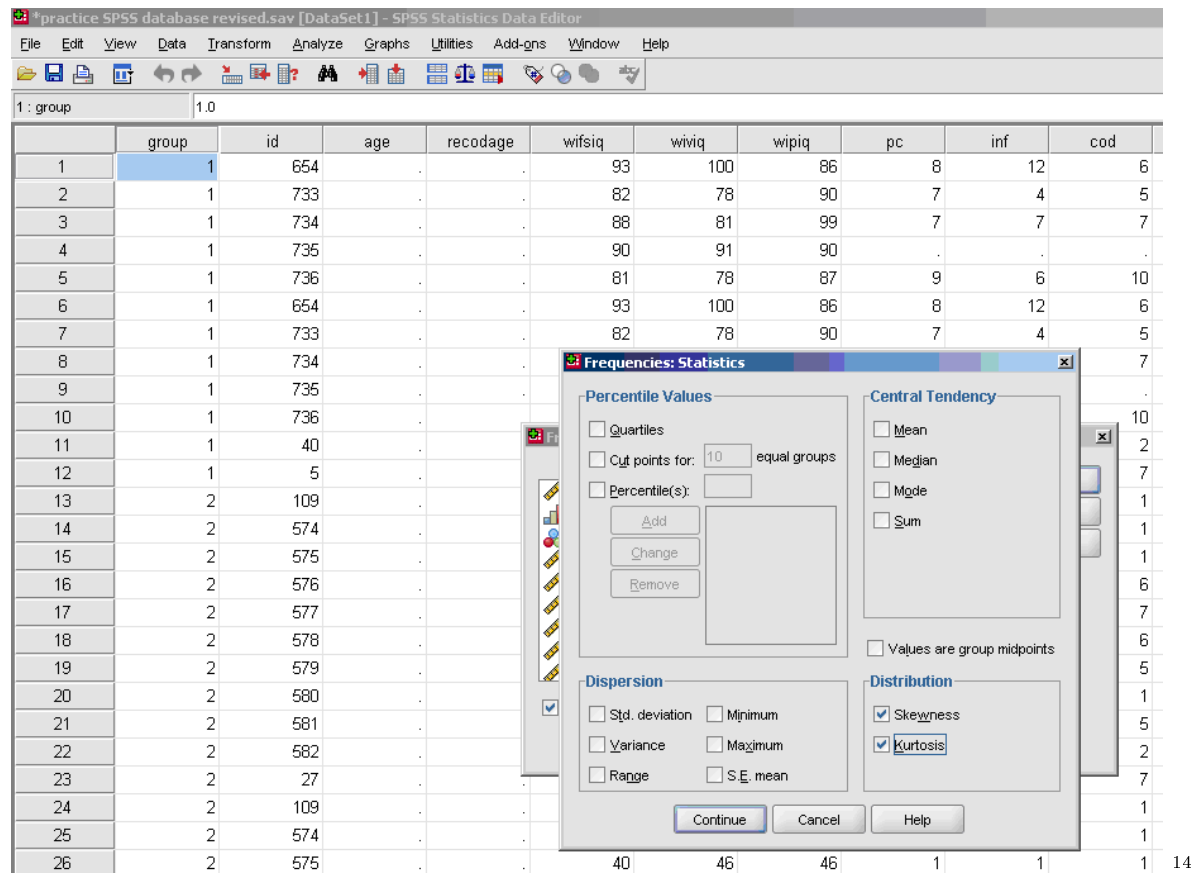
☐ Display frequency tables

OK Paste Reset Cancel Help

13

✓ Click on Statistics
Your screen will now look like this

¹³<http://cnx.org/content/m37334/latest/figure8.5.PNG/image>



* Skewness [Note. Skewness refers to the extent to which the data are normally distributed around the mean. Skewed data involve having either mostly high scores with a few low ones or having mostly low scores with a few high ones.] Readers are referred to the following sources for a more detailed definition of skewness: http://www.statistics.com/index.php?page=glossary&term_id=356 and <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

To standardize the skewness value so that its value can be constant across datasets and across studies, the following calculation must be made: Take the skewness value from the SPSS output and divide it by the Std. error of skewness. If the resulting calculation is within -3 to +3, then the skewness of the dataset is within the range of normality (Onwuegbuzie & Daniel, 2002). If the resulting calculation is outside of this +/-3 range, the dataset is not normally distributed.

* Kurtosis [Note. Kurtosis also refers to the extent to which the data are normally distributed around the mean. This time, the data are piled up higher than normal around the mean or piled up higher than normal at the ends of the distribution.] Readers are referred to the following sources for a more detailed definition of kurtosis: http://www.statistics.com/index.php?page=glossary&term_id=326 and <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

To standardize the kurtosis value so that its value can be constant across datasets and across studies, the following calculation must be made: Take the kurtosis value from the SPSS output and divide it by the Std. error of kurtosis. If the resulting calculation is within -3 to +3, then the kurtosis of the dataset is within the range of normality (Onwuegbuzie & Daniel, 2002). If the resulting calculation is outside of this +/-3 range, the dataset is not normally distributed.

* Continue

* OK

¹⁴<http://cnx.org/content/m37334/latest/figure8.6.PNG/image>

Note: Before you continue to another application you must “UNSPLIT” the files before moving on to other steps:

- ✓ Data
- ✓ Split Files
- ✓ Analyze all cases, do not create groups
- ✓ OK

The screenshot shows the SPSS Statistics Data Editor window with a dataset named 'practice SPSS database revised.sav'. The dataset has 26 rows and 10 columns: group, id, age, recodage, wifsiq, wiviq, wipiq, pc, inf, and cod. A 'Split File' dialog box is open, showing the following options:

- ☐ Analyze all cases, do not create groups
- ☒ Compare groups
- ☐ Organize output by groups

Under 'Groups Based on:', the variable 'Disability Group Membership' is selected. The 'Sort the file by grouping variables' option is also selected. The 'Current Status' at the bottom indicates 'Analysis by groups is off.'.

Check for Skewness and Kurtosis values falling within/without the parameters of normality (-3 to +3). Note that each variable below has its own skewness value and its own kurtosis value. Thus, a total of three standardized skewness coefficients and three standardized kurtosis coefficients can be calculated from information in the table below.

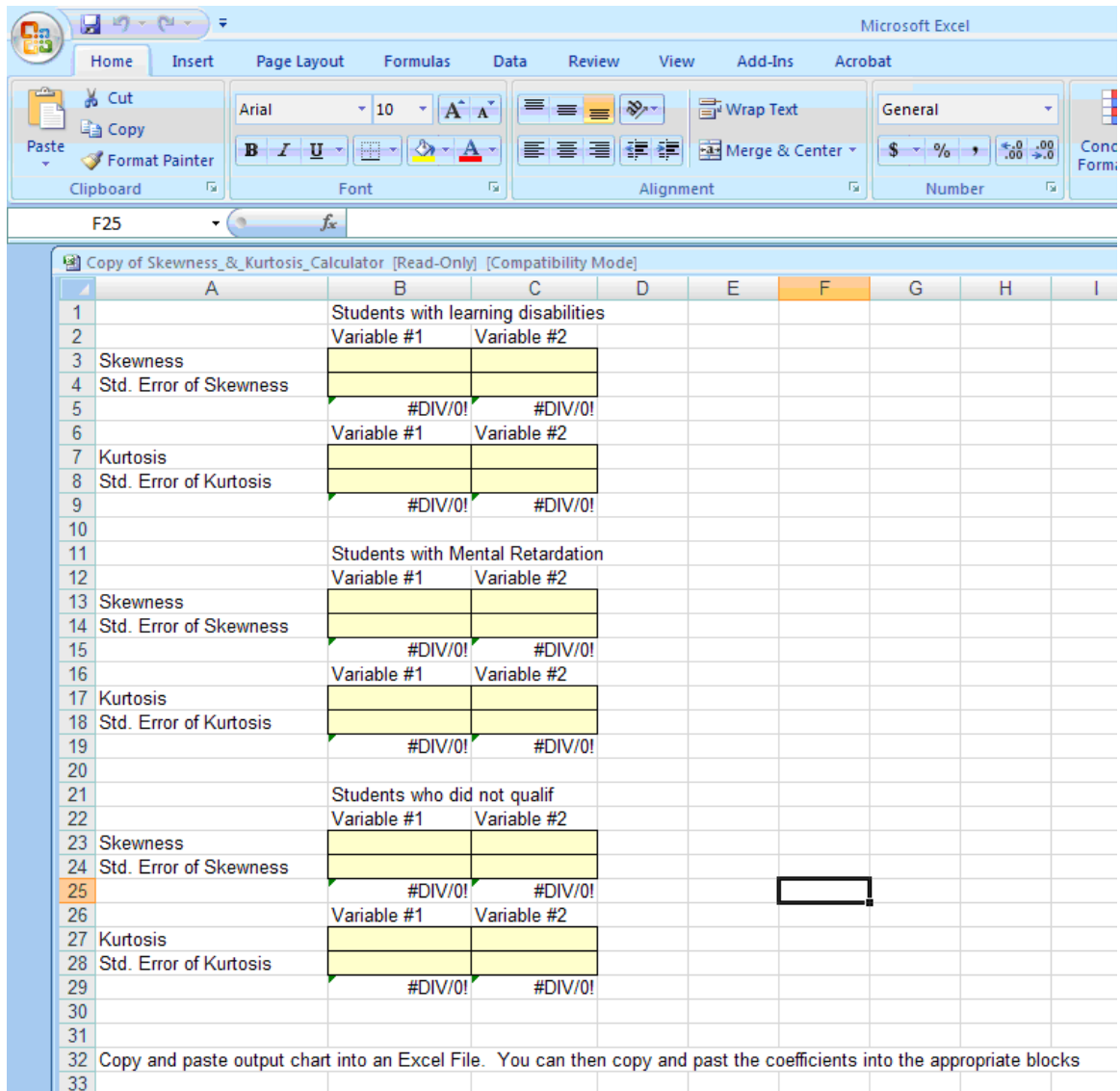
¹⁵<http://cnx.org/content/m37334/latest/figure8.7.PNG/image>

Skewness and Kurtosis Coefficients

	CH005TC09R	CL005TC09R	CW005TC09R
N Valid	3125	1805	1877
Missing	5197	6517	6445
Skewness	-1.129	-.479	-2.197
Std. Error of Skewness	.044	.058	.056
Kurtosis	1.818	-.42	6.991
Std. Error of Kurtosis	.088	.115	.113

Table 9.1

Copy skewness and kurtosis information into the skewness and kurtosis calculator



16

9.2.2 Step Two

Compute Descriptive Statistics on the Dependent Variable

- * Do so via the ANOVA procedure
- * Note. Do not use the ANOVA statistical significance information provided in the output. Use only the *Ms*, *SDs*, and *ns*.
- * **The screen shot will occur in the next step** (Mean and standard deviation)

9.2.3 Step Three

Conduct Analysis of Variance

- ✓ Analyze
- ✓ General Linear Model

¹⁶<http://cnx.org/content/m37334/latest/calc.PNG/image>

✓ Univariate

The screenshot shows the SPSS Statistics Data Editor window for a file named '*practice SPSS database revised.sav [DataSet1]'. The 'Analyze' menu is open, and the path 'General Linear Model' > 'Univariate...' is selected. The data view shows a table with columns: group, id, wifsiq, wiviq, wipiq, and p. The data is organized into rows, with the first 18 rows having a 'group' of 1 and an 'id' ranging from 1 to 18. The last 4 rows (19-22) have a 'group' of 1 and an 'id' of 624, 628, 635, and 664 respectively. The 'wifsiq' column has values 93, 82, 88, 90, 81, 86, 83, 90, 103, 95, 106, 95, 90, 103, 95, 106, 95. The 'wiviq' column has values 100, 78, 81, 91, 78, 88, 85, 91, 98, 99, 98, 93, 91, 98, 99, 98, 93. The 'wipiq' column has values 86, 90, 99, 90, 87, 86, 83, 90, 108, 91, 115, 99, 90, 108, 91, 115, 99. The 'p' column has values 86, 90, 99, 90, 87, 86, 83, 90, 108, 91, 115, 99, 90, 108, 91, 115, 99.

	group	id	wifsiq	wiviq	wipiq	p
1	1					
2	1					
3	1					
4	1					
5	1					
6	1					
7	1					
8	1					
9	1					
10	1					
11	1					
12	1					
13	1					
14	1					
15	1					
16	1					
17	1					
18	1					
19	1	624	6	1.00	103	98
20	1	628	6	1.00	95	99
21	1	635	6	1.00	106	98
22	1	664	6	1.00	95	93

✓ Dependent variable is sent over to the top box, titled dependent variable

✓ Grouping Variable is sent over to the fixed factor box

¹⁷<http://cnx.org/content/m37334/latest/figure8.8.PNG/image>

SPSS Statistics Data Editor window showing a dataset named "practice SPSS database revised.sav [DataSet1]". The dataset contains 25 rows and 11 columns: group, id, age, recodage, wfsiq, wvsiq, wpiq, pc, inf, cod, and an unlabeled column. The data is as follows:

	group	id	age	recodage	wfsiq	wvsiq	wpiq	pc	inf	cod	
1	1	654	.	.	93	100	86	8	12	6	10
2	1	733	.	.	82	78	90	7	4	5	2
3	1	734	.	.	88	81	99	7	7	7	7
4	1	735	.	.	90	91	90	.	.	.	.
5	1	736	.	.	81	78	87	9	6	10	10
6	1	654	.	.	93	100	86	8	12	6	6
7	1	733	.	.	82	78	90	7	4	5	7
8	1	734	.	.	88	81	99	7	7	7	7
9	1	735	.	.	.	.	.	.	.	.	.
10	1	736	.	.	.	.	.	.	.	.	.
11	1	40	.	.	.	.	.	.	.	.	.
12	1	5	.	.	.	.	.	.	.	.	.
13	1	614	6	1.0	.	.	.	.	.	.	10
14	1	624	6	1.0	.	.	.	.	.	.	2
15	1	628	6	1.0	.	.	.	.	.	.	7
16	1	635	6	1.0	.	.	.	.	.	.	9
17	1	664	6	1.0	.	.	.	.	.	.	11
18	1	614	6	1.0	.	.	.	.	.	.	12
19	1	624	6	1.0	.	.	.	.	.	.	19
20	1	628	6	1.0	.	.	.	.	.	.	8
21	1	635	6	1.0	.	.	.	.	.	.	9
22	1	664	6	1.0	.	.	.	.	.	.	11
23	1	522	6	1.0	.	.	.	.	.	.	12
24	1	524	6	1.0	.	.	.	.	.	.	19
25	1	619	6	1.0	.	.	.	.	.	.	8

The "Univariate" dialog box is open, showing the following settings:

- Dependent Variable: Verbal IQ (Wechsler Ve...)
- Fixed Factor(s): Disability Group Membe...
- Random Factor(s):
- Covariate(s):
- WLS Weight:

Buttons: OK, Paste, Reset, Cancel, Help.

- ✓ Options
- ✓ Descriptive Statistics
- ✓ Estimate of effect size
- ✓ Continue

¹⁸<http://cnx.org/content/m37334/latest/figure8.9.PNG/image>

practice SPSS database revised.sav [DataSet1] - SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

1 : group 1.0

	group	id	age	recodage	wfisiq	wiviq	wipiq	pc	inf	cod
101	1	654	.	.	93	100	86	8	12	6
102	1	655	8	1.00	86	84	91	2	5	12
103	1	656	8	1.00	73	73	78	4	6	6
104	1	657	11	3.00	96	104	89	10	11	6
105	1	658	15	4.00	86	78	98	8	5	11
106	1	659	9	2.00	88	84	84	6	6	8
107	1	660	9	2.00	88	84	84	6	6	10
108	1	661	15	4.00	88	84	84	6	6	6
109	1	662	9	2.00	88	84	84	6	6	11
110	1	663	8	1.00	88	84	84	6	6	10
111	1	664	6	1.00	88	84	84	6	6	8
112	1	665	16	4.00	88	84	84	6	6	12
113	1	666	13	4.00	88	84	84	6	6	6
114	1	667	13	4.00	88	84	84	6	6	8
115	1	668	14	4.00	88	84	84	6	6	7
116	1	669	7	1.00	88	84	84	6	6	14
117	1	670	13	4.00	88	84	84	6	6	10
118	1	671	9	2.00	88	84	84	6	6	6
119	1	672	8	1.00	88	84	84	6	6	12
120	1	673	15	4.00	88	84	84	6	6	3
121	1	674	7	1.00	88	84	84	6	6	.
122	1	675	12	3.00	88	84	84	6	6	.
123	1	676	10	2.00	88	84	84	6	6	11
124	1	677	14	4.00	88	84	84	6	6	8
125	1	678	11	3.00	88	84	84	6	6	2
126	1	679	9	2.00	88	84	84	6	6	8
127	1	680	10	2.00	88	84	84	6	6	8

Univariate: Options

Estimated Marginal Means

Factor(s) and Factor Interactions:
(OVERALL)
group

Display Means for:
group

☐ Compare main effects

Confidence interval adjustment:
LSD(none)

Display

☒ Descriptive statistics ☐ Homogeneity tests

☒ Estimates of effect size ☐ Spread vs. level plot

☐ Observed power ☐ Residual plot

☐ Parameter estimates ☐ Lack of fit

☐ Contrast coefficient matrix ☐ General estimable function

Significance level: .05 Confidence intervals are 95.0%

Continue Cancel Help

- ✓ Post Hoc
- ✓ Scheffé
- ✓ Click on variables on which you want the Post Hoc Tests
- ✓ Continue
- ✓ OK

¹⁹<http://cnx.org/content/m37334/latest/overall.PNG/image>

The screenshot shows the SPSS Statistics Data Editor with a dataset named 'practice SPSS database revised.sav [DataSet1]'. The dataset has 29 rows and 11 columns: group, id, age, recodage, wifsiq, wiviq, wipiq, pc, inf, cod. A dialog box titled 'Univariate: Post Hoc Multiple Comparisons for Observed Means' is open, showing 'group' as the factor and 'Scheffe' as the selected test. The dialog also shows options for 'Equal Variances Assumed' and 'Equal Variances Not Assumed'.

9.2.4 Step Four

Check for Statistical Significance

1. Go to the ANOVA table and look at the far right column labeled Sig to check for statistical significance.
2. If you have any value less than .05 then you have statistical significance. Remember to replace the third zero with a 1, if the sig value is .000 (i.e., if the sig value reads as .000, replace the third 0, so it reads as .001).
3. Numerical Sentence = $F(df \text{ between}, df \text{ within})_{sp} = {}_{sp}F \text{ value}_{sp} p_{sp} < {}_{sp}.001$.
4. The outcome of the ANOVA, $F(2,1179) = 503.22, p = .001$, was

Dependent Variable: Verbal IQ (Wechsler Verbal Intelligence 3)

Tests Between-Subjects Effects

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
<i>continued on next page</i>						

²⁰<http://cnx.org/content/m37334/latest/figure8.11.PNG/image>

Corrected	101503.093 ^a	2	50751.547	503.219	.000	.461
Model						
Intercept	6366052.170	1	6366052.170	63121.595	.000	.982
group	101503.093	2	50751.547	503.219	.000	.461
Error	118906.620	1179	100.854			
Total	7405615.000	1182				
Corrected Total	220409.713	1181				

Table 9.2

a. R Squared = .461 (Adjusted R Square = .460)

9.2.5 Step Five

1. Partial Eta² is the effect size n^2

2. Cohen (1988)

.01 - .059 = small effect size

.06 - .139 = moderate effect size

.14 and above = large effect size

Note. n^2 cannot be greater than 1.00. Therefore, a 0 should not be placed in front of the decimal point.

9.2.6 Step Six:

Narrative and Interpretation

1. F value
2. degrees of freedom for groups and for participants
3. p value
4. Post hoc results
5. M , SD , and n for each group (in a table)

9.3 Writing Up Your Statistics

So, how do you "write up" your Research Questions and your Results? Schuler W. Huck (2000) in his seminal book entitled, *Reading Statistics and Research*, points to the importance of your audience understanding and making sense of your research in written form. Huck further states:

9.3.1

This book is designed to help people *decipher* what researchers are trying to communicate in the written or oral summaries of their investigations. Here, the goal is simply to distill meaning from the words, symbols, tables, and figures included in the research report. To be competent in this arena, one must not only be able to decipher what's presented but also to "fill in the holes"; this is the case because researchers typically assume that those receiving the research report are familiar with unmentioned details of the research process and statistical treatment of data.

Researchers and Professors John Slate and Ana Rojas-LeBouef understand this critical issue, so often neglected or not addressed by other authors and researchers. They point to the importance of doctoral students "writing up their statistics" in a way that others can understand your reporting and as importantly,

interpret the meaning of your significant findings and implications for the preparation and practice of educational leadership. Slate and LeBouef provide you with a model for "writing up your parametric ANOVA statistics."

Click here to view: *Writing Up Your Parametric One Way ANOVA Statistics* ²¹

9.4 References

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum

Hyperstats Online Statistics Textbook. (n.d.) Retrieved from <http://davidmlane.com/hyperstat/> ²²

Kurtosis. (n.d.). Definition. Retrieved from http://www.statistics.com/index.php?page=glossary&term_id=326 ²³

Kurtosis. (n.d.). *Definition of normality*. Retrieved from <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb> ²⁴

Onwuegbuzie, A. J., & Daniel, L. G. (2002). Uses and misuses of the correlation coefficient. *Research in the Schools*, 9(1), 73-90.

Skewness. (n.d.) Retrieved from http://www.statistics.com/index.php?page=glossary&term_id=356 ²⁵

Skewness. (n.d.). *Definition of normality*. Retrieved from <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb> ²⁶

StatSoft, Inc. (2011). *Electronic statistics textbook*. Tulsa, OK: StatSoft. WEB: <http://www.statsoft.com/textbook/> ²⁷

²¹<http://cnx.org/content/col11299/latest/>

²²<http://davidmlane.com/hyperstat/>

²³http://www.statistics.com/index.php?page=glossary&term_id=326

²⁴<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

²⁵http://www.statistics.com/index.php?page=glossary&term_id=356

²⁶<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

²⁷<http://www.statsoft.com/textbook/>

Chapter 10

Conducting a Nonparametric One-Way Analysis of Variance



NOTE: This chapter has been peer-reviewed, accepted, and endorsed by the National Council of Professors of Educational Administration (NCPEA) as a significant contribution to the scholarship and practice of education administration. Formatted and edited in Connexions by Theodore Creighton and Brad Bizzell, Virginia Tech, Janet Tareilo, Stephen F. Austin State University, and Thomas Kersten, Roosevelt University.

10.1

This chapter is part of a larger Collection (Book) and is available at: *Calculating Basic Statistical Procedures in SPSS: A Self-Help and Practical Guide to Preparing Theses, Dissertations, and Manuscripts* ²

NOTE: Slate and LeBouef have written a "companion book" which is available at: *Preparing and Presenting Your Statistical Findings: Model Write Ups* ³

Authors Information

John R. Slate is a Professor at Sam Houston State University where he teaches Basic and Advanced Statistics courses, as well as professional writing, to doctoral students in Educational Leadership and Counseling. His research interests lie in the use of educational databases, both state and national, to reform school practices. To date, he has chaired and/or served over 100 doctoral student dissertation committees. Recently, Dr. Slate created a website Writing and Statistical Help⁴ to assist students and faculty with both statistical assistance and in editing/writing their dissertations/theses and manuscripts.

¹This content is available online at <<http://cnx.org/content/m37335/1.6/>>.

²<http://my.qoop.com/store/NCPEA-Publications-1781472103076212/>

³<http://my.qoop.com/store/NCPEA-Publications-1781472103076212/>

⁴<http://www.writingandstatisticalhelp.com>

Ana Rojas-LeBouef is a Literacy Specialist at the Reading Center at Sam Houston State University where she teaches developmental reading courses. She recently completed her doctoral degree in Reading, where she conducted a 16-year analysis of Texas statewide data regarding the achievement gap. Her research interests lie in examining the inequities in achievement among ethnic groups. Dr. Rojas-LeBouef also assists students and faculty in their writing and statistical needs on the website Writing and Statistical Help.⁵

Editors Information

Theodore B. Creighton, is a Professor at Virginia Tech and the Publications Director for NCPEA Publications⁶, the Founding Editor of Education Leadership Review,⁷ and the Senior Editor of the NCPEA Connexions Project.

Brad E. Bizzell, is a recent graduate of the Virginia Tech Doctoral Program in Educational Leadership and Policy Studies, and is a School Improvement Coordinator for the Virginia Tech Training and Technical Assistance Center. In addition, Dr. Bizzell serves as an Assistant Editor of the NCPEA Connexions Project in charge of technical formatting and design.

Janet Tareilo, is a Professor at Stephen F. Austin State University and serves as the Assistant Director of NCPEA Publications. Dr. Tareilo also serves as an Assistant Editor of the NCPEA Connexions Project and as a editor and reviewer for several national and international journals in educational leadership.

Thomas Kersten is a Professor at Roosevelt University in Chicago. Dr. Kersten is widely published and an experienced editor and is the author of Taking the Mystery Out of Illinois School Finance⁸, a Connexions Print on Demand publication. He is also serving as Editor in Residence for this book by Slate and LeBouef.

10.2 Nonparametric One-Way Analysis of Variance

In this set of steps, readers will calculate either a parametric or a nonparametric statistical analysis, depending on whether the data for the dependent variable reflect a normal distribution. A parametric statistical procedure requires that its data be reflective of a normal curve whereas no such assumption is made in the use of a nonparametric procedure. Of the two types of statistical analyses, the parametric procedure is the more powerful one in ascertaining whether or not a statistically significant difference, in this case, exists. As such, parametric procedures are preferred over nonparametric procedures. When data are not normally distributed, however, parametric analyses may provide misleading and inaccurate results. According, nonparametric analyses should be used in cases where data are not reflective of a normal curve. In this set of steps, readers are provided with information on how to make the determination of normally or nonnormally distributed data. For detailed information regarding the assumptions underlying parametric and nonparametric procedures, readers are referred to the Hyperstats Online Statistics Textbook at <http://davidmlane.com/hyperstat/> or to the Electronic Statistics Textbook (2011) at <http://www.statsoft.com/textbook/>

For this nonparametric analysis of variance procedure to be appropriately used, at least half of the standardized skewness coefficients and the standardized kurtosis coefficients must be outside the normal range (+/-3, Onwuegbuzie & Daniel, 2002). Research questions for which nonparametric analysis of variance procedures are appropriate involve asking for differences in a dependent variable by group membership (i.e., more than two groups may be present). The research question, “What is the difference in science performance among middle school students as a function of ethnic membership?” could be answered through use of a nonparametric analysis of variance procedure.

⁵<http://www.writingandstatisticalhelp.com>

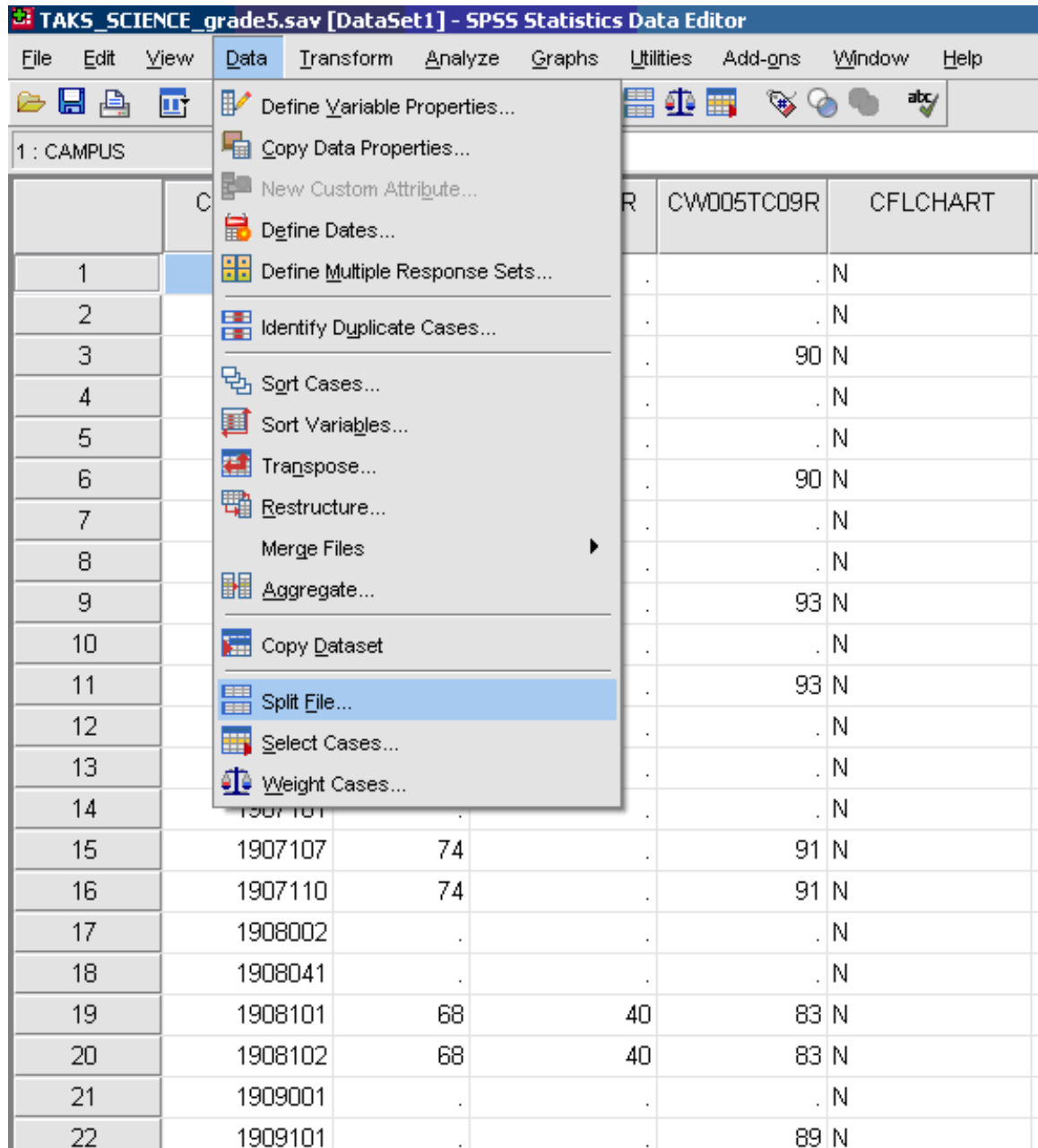
⁶<http://www.ncpeapublications.org>

⁷<http://ncpeapublications.org/about-elr.html>

⁸<http://cnx.org/content/col10606/latest/>

10.2.1 Step One:

✓ Split your file on the basis on your independent variable/fixed factor/grouping variable



After you do this, your screen should resemble the following:

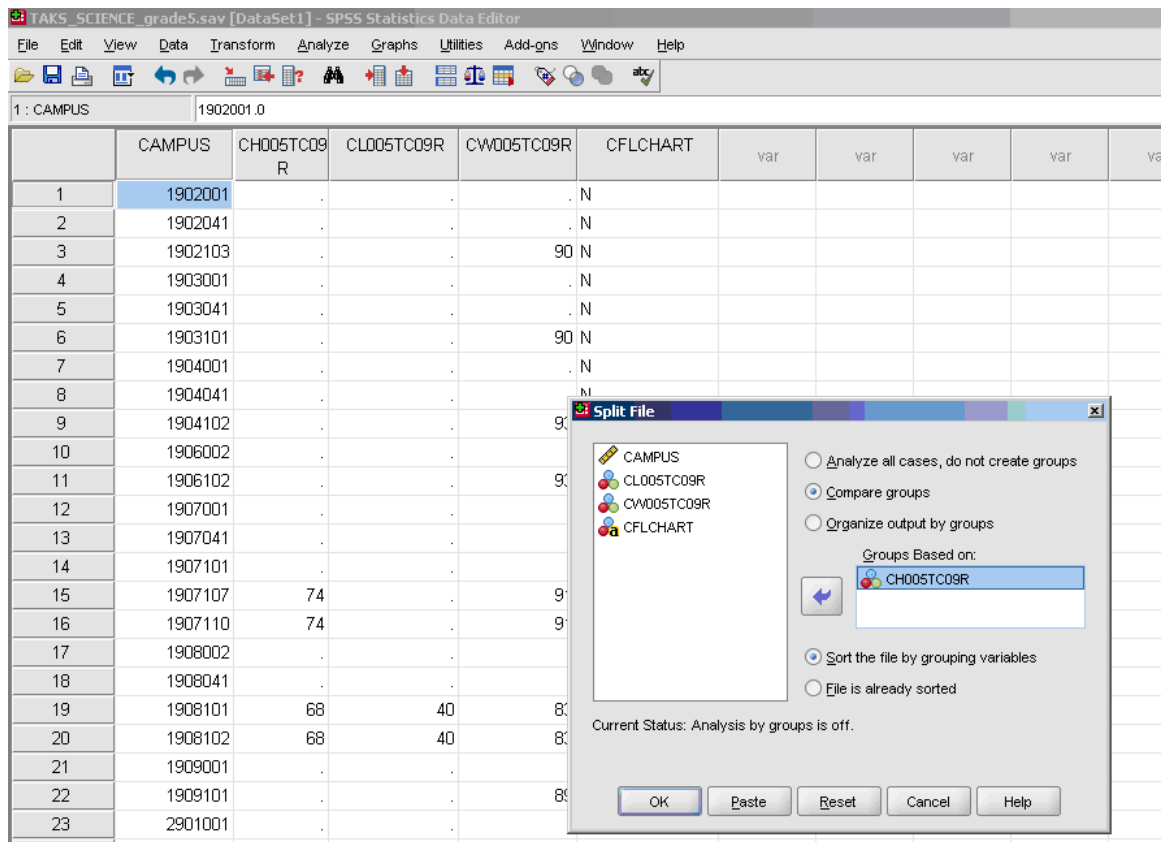
⁹<http://cnx.org/content/m37335/latest/figure9.1.PNG/image>

The screenshot shows the SPSS Statistics Data Editor window with the file 'TAKS SCIENCE grade5.sav [DataSet1]'. The dataset has 24 rows and 10 columns. The 'CAMPUS' column is highlighted. A 'Split File' dialog box is open, showing the 'CAMPUS' variable selected for grouping. The 'Analyze all cases, do not create groups' option is selected.

	CAMPUS	CH005TC09R	CL005TC09R	CW005TC09R	CFLCHART	var	var	var	var
1	1902001	.	.	.	N				
2	1902041	.	.	.	N				
3	1902103	.	.	90	N				
4	1903001	.	.	.	N				
5	1903041	.	.	.	N				
6	1903101	.	.	90	N				
7	1904001	.	.	.	N				
8	1904041	.	.	.	N				
9	1904102	.	.	90	N				
10	1906002	.	.	.	N				
11	1906102	.	.	90	N				
12	1907001	.	.	.	N				
13	1907041	.	.	.	N				
14	1907101	.	.	.	N				
15	1907107	74	.	90	N				
16	1907110	74	.	90	N				
17	1908002	.	.	.	N				
18	1908041	.	.	.	N				
19	1908101	68	40	80	N				
20	1908102	68	40	80	N				
21	1909001	.	.	.	N				
22	1909101	.	.	80	N				
23	2901001	.	.	.	N				
24	2901002	.	.	.	N				

Your screen will show that all cases are going to be analyzed and a “do not create groups”. You will need to click the compare groups and move the dependent variable over to the “Group Based on”.

¹⁰<http://cnx.org/content/m37335/latest/figure9.2.PNG/image>



11

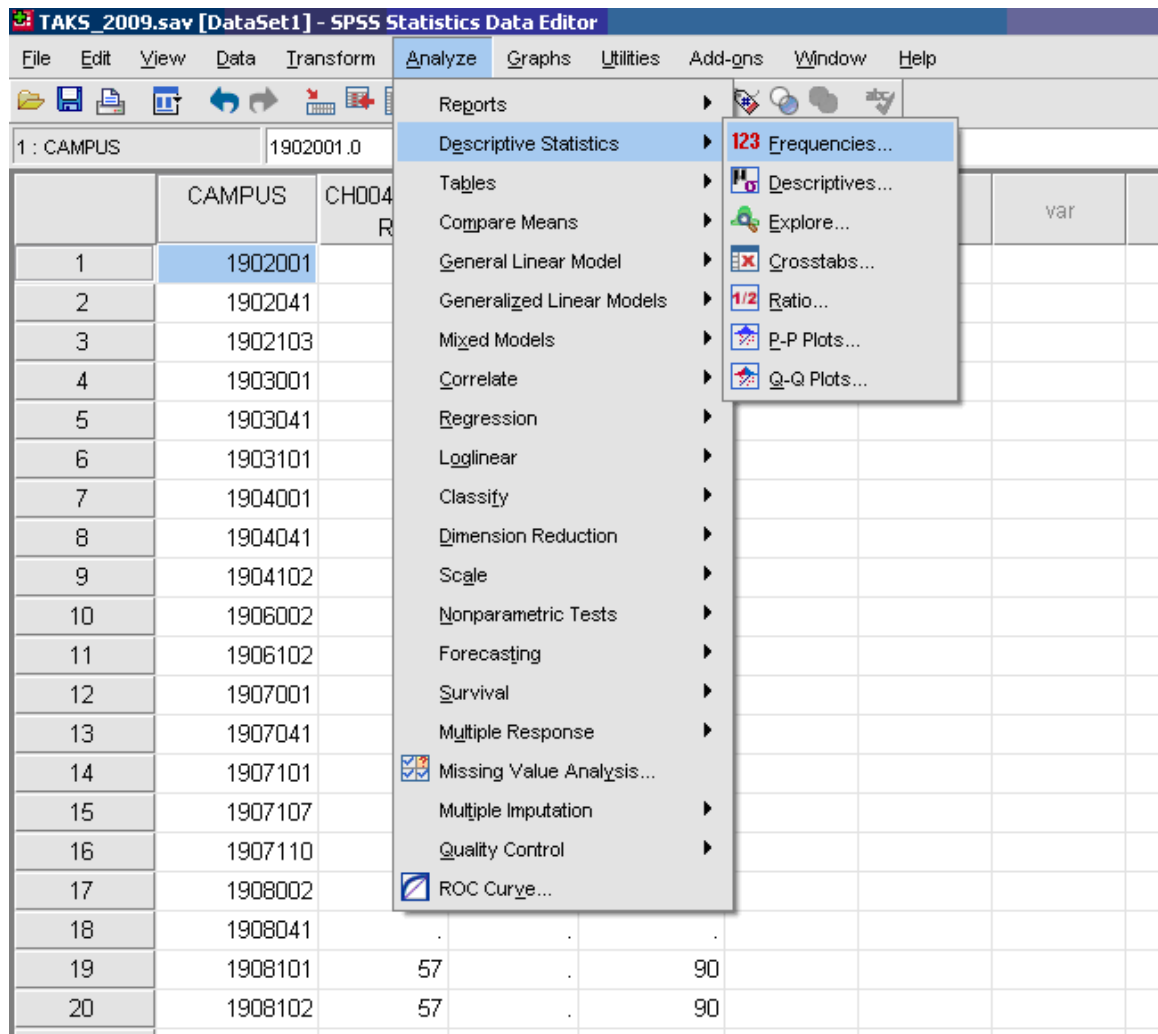
* Click OK

✓ Analyze

* Descriptive Statistics

* Frequencies

¹¹<http://cnx.org/content/m37335/latest/figure9.3.PNG/image>



12

✓ Move over the dependent (outcome) variable

¹²<http://cnx.org/content/m37335/latest/figure9.4.PNG/image>

TAKS_SCIENCE_grade5.sav [DataSet1] - SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

1 : CAMPUS 1902001.0

	CAMPUS	CH005TC09R	CL005TC09R	CW005TC09R	CFLCHART	var	var	var
1	1902001	.	.	.	N			
2	1902041	.	.	.	N			
3	1902103	.	.	90	N			
4	1903001	.	.	.	N			
5	1903041	.	.	.	N			
6	1903101	.	.	90	N			
7	1904001	.	.	.	N			
8	1904041	.	.	.	N			
9	1904102	.	.	93	N			
10	1906002	.	.	.	N			
11	1906102	.	.	.				
12	1907001	.	.	.				
13	1907041	.	.	.				
14	1907101	.	.	.				
15	1907107	74	.	.				
16	1907110	74	.	.				
17	1908002	.	.	.				
18	1908041	.	.	.				
19	1908101	68	40	.				
20	1908102	68	40	.				
21	1909001	.	.	.				
22	1909101	.	.	.				
23	2901001	.	.	.				

Frequencies

Variable(s): CH005TC09R

Statistics...
Charts...
Format...

☐ Display frequency tables

OK Paste Reset Cancel Help

13

✓ Click on Statistics
Your screen will look like this.

¹³<http://cnx.org/content/m37335/latest/outcome.PNG/image>

The screenshot shows the SPSS Statistics Data Editor window with a dataset named 'practice SPSS database revised.sav'. The dataset has 25 rows and 11 columns: group, id, age, recodage, wifsiq, wiviq, wipiq, pc, inf, cod. A 'Frequencies: Statistics' dialog box is open, showing options for Percentile Values, Central Tendency, Dispersion, and Distribution. The 'Distribution' section has 'Skewness' and 'Kurtosis' checked.

	group	id	age	recodage	wifsiq	wiviq	wipiq	pc	inf	cod
1	1	654	.	.	93	100	86	8	12	.
2	1	733	.	.	82	78	90	7	4	.
3	1	734	.	.	88	81	99	7	7	.
4	1	735	.	.	90	91	90	.	.	.
5	1	736	.	.	81	78	87	9	6	.
6	1	654	.	.	93	100	86	8	12	.
7	1	733	.	.	82	78	90	7	4	.
8	1	734	.	.	.	.	.	.	.	.
9	1	735	.	.	.	.	.	.	.	.
10	1	736	.	.	.	.	.	.	.	.
11	1	40	.	.	.	.	.	.	.	.
12	1	5	.	.	.	.	.	.	.	.
13	2	109	.	.	.	.	.	.	.	.
14	2	574	.	.	.	.	.	.	.	.
15	2	575	.	.	.	.	.	.	.	.
16	2	576	.	.	.	.	.	.	.	.
17	2	577	.	.	.	.	.	.	.	.
18	2	578	.	.	.	.	.	.	.	.
19	2	579	.	.	.	.	.	.	.	.
20	2	580	.	.	.	.	.	.	.	.
21	2	581	.	.	.	.	.	.	.	.
22	2	582	.	.	.	.	.	.	.	.
23	2	27	.	.	.	.	.	.	.	.
24	2	109	.	.	.	.	.	.	.	.
25	2	574	.	.	.	.	.	.	.	.

14

* Skewness [Note. Skewness refers to the extent to which the data are normally distributed around the mean. Skewed data involve having either mostly high scores with a few low ones or having mostly low scores with a few high ones.] Readers are referred to the following sources for a more detailed definition of skewness: http://www.statistics.com/index.php?page=glossary&term_id=356 and <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

To standardize the skewness value so that its value can be constant across datasets and across studies, the following calculation must be made: Take the skewness value from the SPSS output and divide it by the Std. error of skewness. If the resulting calculation is within -3 to +3, then the skewness of the dataset is within the range of normality (Onwuegbuzie & Daniel, 2002). If the resulting calculation is outside of this +/-3 range, the dataset is not normally distributed.

* Kurtosis [Note. Kurtosis also refers to the extent to which the data are normally distributed around the mean. This time, the data are piled up higher than normal around the mean or piled up higher than normal at the ends of the distribution.] Readers are referred to the following sources for a more detailed definition of kurtosis: http://www.statistics.com/index.php?page=glossary&term_id=326 and <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

To standardize the kurtosis value so that its value can be constant across datasets and across studies, the following calculation must be made: Take the kurtosis value from the SPSS output and divide it by the Std. error of kurtosis. If the resulting calculation is within -3 to +3, then the kurtosis of the dataset is within the range of normality (Onwuegbuzie & Daniel, 2002). If the resulting calculation is outside of this +/-3 range, the dataset is not normally distributed.

* Continue

* OK

¹⁴<http://cnx.org/content/m37335/latest/figure9.6.PNG/image>

Note: Before you continue to another application you must “UNSPLIT” the files before moving on to other steps:

- ✓ Data
- ✓ Split Files
- ✓ Analyze all cases, do not create groups
- ✓ OK

The screenshot shows the SPSS Statistics Data Editor window with the file 'TAKS_SCIENCE_grade5.sav [DataSet1]'. The data is organized into a table with 25 rows and 10 columns. The first column is 'CAMPUS', and the subsequent columns are 'CH005TC09R', 'CL005TC09R', 'CW005TC09R', 'CFLCHART', and four empty columns labeled 'var'. The 'Split File' dialog box is open, showing the 'CAMPUS' variable selected for splitting. The 'Analyze all cases, do not create groups' option is selected. The 'Current Status' is 'Compare: CH005TC09R'.

	CAMPUS	CH005TC09R	CL005TC09R	CW005TC09R	CFLCHART	var	var	var	var
1	1902001	.	.	.	N				
2	1902041	.	.	.	N				
3	1902103	.	.	90	N				
4	1903001	.	.	.	N				
5	1903041	.	.	.	N				
6	1903101	.	.	90	N				
7	1904001	.	.	.	N				
8	1904041	.	.	.	N				
9	1904102	.	.	90	N				
10	1906002	.	.	.	N				
11	1906102	.	.	90	N				
12	1907001	.	.	.	N				
13	1907041	.	.	.	N				
14	1907101	.	.	.	N				
15	1908002	.	.	.	N				
16	1908041	.	.	.	N				
17	1909001	.	.	.	N				
18	1909101	.	.	80	N				
19	2901001	.	.	.	N				
20	2901002	.	.	.	N				
21	2901041	.	.	.	N				
22	2901101	.	.	.	N				
23	2901105	.	.	.	N				
24	3801001	.	.	.	N				
25	3902001	.	.	.	N				

Check for Skewness and Kurtosis values falling within/without the parameters of normality (-3 to +3). Note that each variable below has its own skewness value and its own kurtosis value. Thus, a total of three standardized skewness coefficients and three standardized kurtosis coefficients can be calculated from information in the table below.

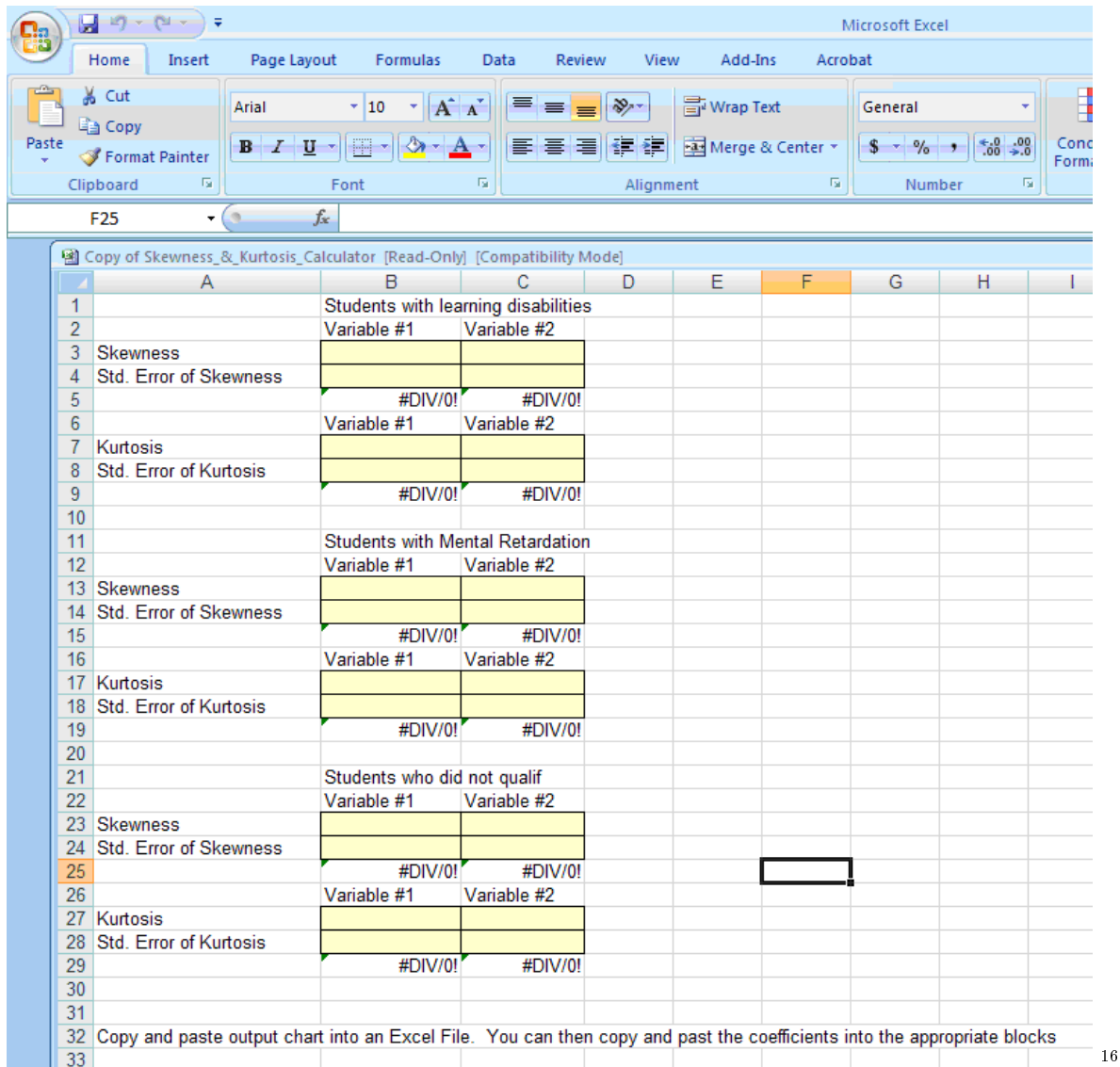
¹⁵<http://cnx.org/content/m37335/latest/figure9.7.PNG/image>

Skewness and Kurtosis Coefficients

	CH005TC09R	CL005TC09R	CW005TC09R
N Valid	3125	1805	1877
Missing	5197	6517	6445
Skewness	-1.129	-.479	-2.197
Std. Error of Skewness)	.044	.058	.056
Kurtosis	1.818	-.412	6.991
Std. Error of Kurtosis	.008	.115	.113

Table 10.1

Copy skewness and kurtosis information into the skewness and kurtosis calculator



10.2.2 Step Two:

Compute Descriptive Statistics on the Dependent Variable

- * Do so via the ANOVA procedure.
- * Note. Do not use the ANOVA statistical significance information provided in the output. Use only the *Ms*, *SDs*, and *ns*.
- * **The screen shot will occur in the next step** (Mean and standard deviation)

10.2.3 Step Three:

Run Nonparametric One-Way ANOVA on Data

- * Analyze
- * Nonparametric Tests

¹⁶<http://cnx.org/content/m37335/latest/calc.PNG/image>

* k Independent Samples

The screenshot shows the SPSS Statistics Data Editor window with the file name '*practice SPSS database revised.sav [DataSet1]'. The 'Analyze' menu is open, and the path 'Nonparametric Tests' > 'K Independent Samples...' is highlighted. The data view shows a table with columns 'group', 'id', 'wifsiq', 'wiviq', and 'wipiq'. The 'group' column has values 1 and 2, and the 'id' column has values from 1 to 22. The 'wifsiq', 'wiviq', and 'wipiq' columns contain numerical data for each group.

group	id	wifsiq	wiviq	wipiq
1	1	93	100	86
1	2	82	78	90
1	3	88	81	99
1	4	90	91	90
1	5	81	78	87
1	6	93	100	86
1	7	82	78	90
1	8	88	81	99
1	9	90	91	90
1	10	81	78	87
1	11			
1	12			
2	13			
2	14			
2	15			
2	16			
2	17			
2	18			
2	19	579	64	55
2	20	580	44	46
2	21	581	57	66
2	22	582	47	55

17

* Keep the default of Kruskal-Wallis H checked

¹⁷<http://cnx.org/content/m37335/latest/figure9.8.PNG/image>

1 : group 1.0

	group	id	age	recodage	wfsiq	wviq	wpiq	pc	inf	cod
1	1	110	14	4.00	83	81	89	10	7	8
2	1	111	12	3.00	80	82	81	8	8	7
3	1	112	15	4.00	90	91	90	10	8	2
4	1	122	8	1.00	73	71	80	3	4	11
5	1	123	7	1.00	80	76	87	5	7	17
6	1	124	9	2.00	76	75	80	2	5	8
7	1	131	13	4.00	82	78	90	7	4	5
8	1	133	12	3.00	88	81	99	7	7	7
9	1	137	9	2.00	84	73	87	6	6	10
10	1	201	8	1.00	81	79	84	7	7	7
11	1	202	8	1.00	81	79	84	7	7	5
12	1	203	13	4.00	81	79	84	7	7	5
13	1	206	10	2.00	81	79	84	7	7	11
14	1	208	9	2.00	81	79	84	7	7	8
15	1	245	9	2.00	81	79	84	7	7	3
16	1	246	11	3.00	81	79	84	7	7	5
17	1	249	9	2.00	81	79	84	7	7	7
18	1	254	15	4.00	81	79	84	7	7	5
19	1	256	14	4.00	81	79	84	7	7	4
20	1	257	11	3.00	81	79	84	7	7	10
21	1	258	12	3.00	81	79	84	7	7	8
22	1	275	13	4.00	81	79	84	7	7	12
23	1	278	7	1.00	81	79	84	7	7	10
24	1	279	13	4.00	81	79	84	7	7	6
25	1	280	16	4.00	86	84	84	4	7	6
26	1	281	13	4.00	69	69	74	5	5	8
27	1	283	9	2.00	73	73	74	4	4	4
28	1	284	14	4.00	80	80	74	7	8	8
29	1	285	9	2.00	73	73	74	4	4	4
30	1	286	12	3.00	82	82	74	5	4	4
31	1	287	10	2.00	84	84	74	4	9	9
32	1	584	10	2.00	75	80	74	5	9	6
33	1	585	15	4.00	92	91	94	11	9	8

18

- * Test Variable would be your Dependent Variable (e.g., test scores)
- * Grouping Variable would be your Independent Variable (categories)
- * Define Groups
- * Insert the number for your lowest numbered group and then the number for your highest numbered group.

¹⁸<http://cnx.org/content/m37335/latest/figure9.10PNG/image>

The screenshot shows the SPSS Statistics Data Editor window with a dataset named 'practice SPSS database revised.sav'. The dataset has 24 rows and 11 columns: group, id, age, recodage, wfsiq, wiviq, wipiq, pc, inf, and cod. The 'group' column has two distinct values: 1 and 2. A dialog box titled 'Tests for Several Independent Samples' is open, showing 'id' as the Test Variable List and 'group(1 2)' as the Grouping Variable. The 'Kruskal-Wallis H' test is selected under the 'Test Type' section.

	group	id	age	recodage	wfsiq	wiviq	wipiq	pc	inf	cod
1	1	654	.	.	93	100	86	8	12	
2	1	733	.	.	82	78	90	7	4	
3	1	734	.	.	88	81	99	7	7	
4	1	735	.	.	90	91	90	.	.	
5	1	736	.	.	81	78	87	9	6	
6	1	654	.	.	93	100	86	8	12	
7	1	733	.	.	82	78	90	7	4	
8	1	734	.	.	88	81	99	7	7	
9	1	735	.	.	90	91	90	.	.	
10	1	736	.	.	81	78	87	9	6	
11	1	40	.	.						
12	1	5	.	.						
13	2	109	.	.						
14	2	574	.	.						
15	2	575	.	.						
16	2	576	.	.						
17	2	577	.	.						
18	2	578	.	.						
19	2	579	.	.						
20	2	580	.	.						
21	2	581	.	.						
22	2	582	.	.						
23	2	27	.	.						
24	2	109	.	.						

19

Note: Click on view than value labels to find the code for each group.

* Continue

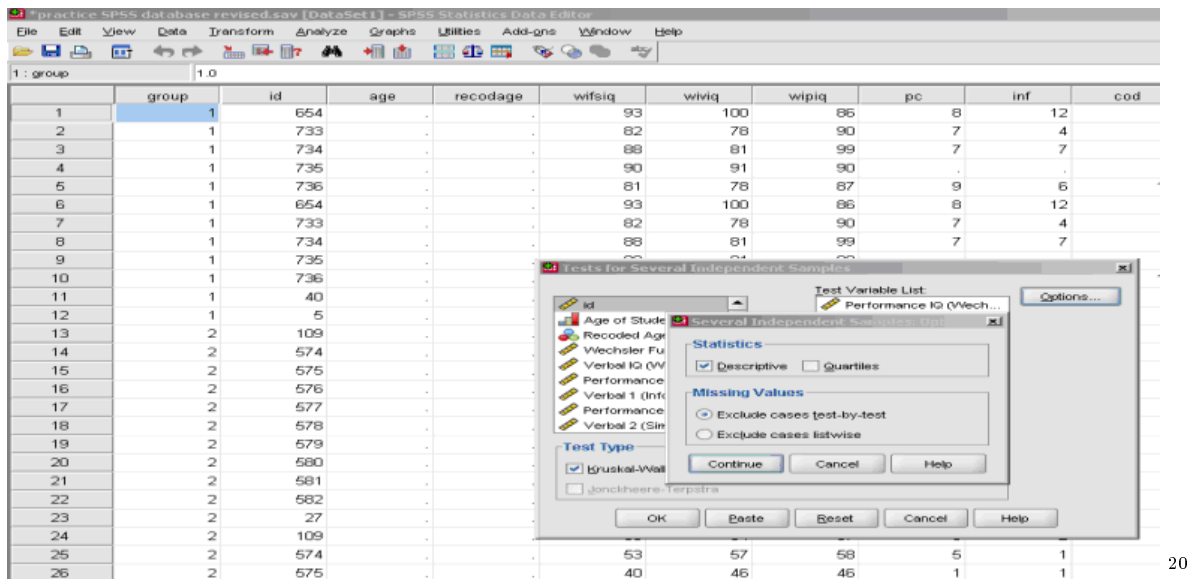
** To obtain the Means and Standard Deviation:

* Click on options

* Highlight Descriptive

* Click Continue

¹⁹<http://cnx.org/content/m37335/latest/figure9.9.PNG/image>



* OK

10.2.4 Step Four:

Check for Statistical Significance

Numerical sentence is written as: $X^2 = 430.66, p < .001$

Test Statistics^{a,b}

	Performance IQ (Wechsler Performance Intelligence 3)
Chi-Square	430.661
df	1
Asymp. Sig.	.000

Table 10.2

a. Kruskal Wallis Test

b. Grouping Variable: Disability Group Membership

10.2.5 Step Five:

If you have a statistically significant finding in your nonparametric ANOVA, you need to run the appropriate nonparametric post hocs. Refer to your steps on running the nonparametric independent samples *t*-test.

²⁰<http://cnx.org/content/m37335/latest/figure9.11.PNG/image>

10.2.6 Step Six:**Calculate Nonparametric Independent Samples t -test on Data**

✓ Analyze

✓ Nonparametric Tests

✓ 2 Independent Samples

✓ Test Variable would be your Dependent Variable (e.g., test scores)

✓ Grouping Variable would be your dichotomous Independent Variable

✓ Define Groups

✓ Group One is No. 1 and Group Two is No. 2 (or whatever numbers you used to identify each group)

Note: Click on view then value labels to find the code for each group

✓ Continue

✓ OK

Note. The above procedure is repeated for each pairwise comparison. Thus, if you have three groups, you would have three calculations. Correct for inflated error by using the Bonferroni method of adjustment. Take the number of pairwise comparisons you are calculating and divide .05 by that.

Check for Statistical Significance

1. Go to the Test Statistics Box and look at the cell in the bottom right column to check for statistical significance.

If you have any value less than .05 then you have statistical significance. Remember to replace the third zero with a 1, if the value is .000 (i.e., for a sig value of .000, thus it would read .001).

Test Statistics^a

	Performance IQ (Wechsler Performance Intelligence 3)
Mann-Whitney U	6765.500
Wilcoxon W	44166.500
Z	-20.752
Asymp. Sig. (2-tailed)	.000

Table 10.3

a. Grouping Variable: Disability Group Membership

To determine how to report the results of these nonparametric followup procedures, see the chapter on nonparametric independent samples t -test in this book.

10.3 Writing Up Your Nonparametric ANOVA

So, how do you "write up" your Research Questions and your Results? Schuler W. Huck (2000) in his seminal book entitled, *Reading Statistics and Research*, points to the importance of your audience understanding and making sense of your research in written form. Huck further states:

10.3.1

This book is designed to help people *decipher* what researchers are trying to communicate in the written or oral summaries of their investigations. Here, the goal is simply to distill meaning from the words, symbols, tables, and figures included in the research report. To be competent in this arena, one must not only be

able to decipher what's presented but also to "fill in the holes"; this is the case because researchers typically assume that those receiving the research report are familiar with unmentioned details of the research process and statistical treatment of data.

A Note from the Editors

Researchers and Professors John Slate and Ana Rojas-LeBouef understand this critical issue, so often neglected or not addressed by other authors and researchers. They point to the importance of doctoral students "writing up their statistics" in a way that others can understand your reporting and as importantly, interpret the meaning of your significant findings and implications for the preparation and practice of educational leadership. Slate and LeBouef provide you with a model for "writing up your Nonparametric ANOVA statistics."

Click here to view: *Writing Up Your Nonparametric ANOVA Statistics* ²¹

10.4 References

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum

Hyperstats Online Statistics Textbook. (n.d.) Retrieved from <http://davidmlane.com/hyperstat/> ²²

Kurtosis. (n.d.). Definition. Retrieved from http://www.statistics.com/index.php?page=glossary&term_id=326 ²³

Kurtosis. (n.d.). *Definition of normality*. Retrieved from <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb> ²⁴

Onwuegbuzie, A. J., & Daniel, L. G. (2002). Uses and misuses of the correlation coefficient. *Research in the Schools*, 9(1), 73-90.

Skewness. (n.d.) Retrieved from http://www.statistics.com/index.php?page=glossary&term_id=356 ²⁵

Skewness. (n.d.). *Definition of normality*. Retrieved from <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb> ²⁶

StatSoft, Inc. (2011). *Electronic statistics textbook*. Tulsa, OK: StatSoft. WEB: <http://www.statsoft.com/textbook/> ²⁷

²¹<http://cnx.org/content/col11299/latest/>

²²<http://davidmlane.com/hyperstat/>

²³http://www.statistics.com/index.php?page=glossary&term_id=326

²⁴<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

²⁵http://www.statistics.com/index.php?page=glossary&term_id=356

²⁶<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

²⁷<http://www.statsoft.com/textbook/>

Chapter 11

Standardized Skewness and Standardized Kurtosis Coefficient Calculator

11.1 Skewness and Kurtosis Calculator

The Standardized Skewness and Kurtosis Coefficient Calculator is to be used with the Book, *Calculating Basic Statistical Procedures in SPSS: A Self-Help and Practical Guide to Preparing Theses, Dissertations, and Manuscripts* authored by Professors John R. Slate and Ana Rojas-LeBouef.

Do not try to OPEN the file from here. DOWNLOAD the Excel file to your desktop, where you can use the Calculator to create your Skewness and Kurtosis coefficients.

Click here to download the Skewness and Kurtosis Excel file to your desktop.²

¹This content is available online at <<http://cnx.org/content/m37347/1.3/>>.

²See the file at

<http://cnx.org/content/m37347/latest/Standardized_Skewness_and_Standardized_Kurtosis_Coefficient_Calculator.xls>

Available for free at Connexions <<http://cnx.org/content/col11292/1.6>>

Chapter 12

Resources: Calculating Basic Statistics in SPSS¹

This chapter is part of a larger Collection (Book) and is available at: *Calculating Basic Statistical Procedures in SPSS: A Self-Help and Practical Guide to Preparing Theses, Dissertations, and Manuscripts* ²

NOTE: Slate and LeBouef have written a "companion book" which is available at: *Preparing and Presenting Your Statistical Findings: Model Write Ups* ³



NOTE: These Recommended Resources have been peer-reviewed, accepted, and endorsed by the National Council of Professors of Educational Administration (NCPEA) as a significant contribution to the scholarship and practice of education administration. Formatted and edited in Connexions by Theodore Creighton and Brad Bizzell, Virginia Tech, Janet Tareilo, Stephen F. Austin State University, and Thomas Kersten, Roosevelt University.

12.1 Recommended Resources

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum

Hyperstats Online Statistics Textbook. (n.d.) Retrieved from <http://davidmlane.com/hyperstat/> ⁴

Kurtosis. (n.d.). Definition. Retrieved from http://www.statistics.com/index.php?page=glossary&term_id=326 ⁵

Kurtosis. (n.d.). *Definition of normality*. Retrieved from <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb> ⁶

Onwuegbuzie, A. J., & Daniel, L. G. (2002). Uses and misuses of the correlation coefficient. *Research in the Schools*, 9(1), 73-90.

¹This content is available online at <<http://cnx.org/content/m37336/1.3/>>.

²<http://my.qoop.com/store/NCPEA-Publications-1781472103076212/>

³<http://my.qoop.com/store/NCPEA-Publications-1781472103076212/>

⁴<http://davidmlane.com/hyperstat/>

⁵http://www.statistics.com/index.php?page=glossary&term_id=326

⁶<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

Skewness. (n.d.) Retrieved from http://www.statistics.com/index.php?page=glossary&term_id=356⁷

Skewness. (n.d.). *Definition of normality*. Retrieved from <http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>⁸

StatSoft, Inc. (2011). *Electronic statistics textbook*. Tulsa, OK: StatSoft. WEB: <http://www.statsoft.com/textbook/>⁹

⁷http://www.statistics.com/index.php?page=glossary&term_id=356

⁸<http://www.statsoft.com/textbook/basic-statistics/#Descriptive%20statisticsb>

⁹<http://www.statsoft.com/textbook/>

Attributions

Collection: *Calculating Basic Statistical Procedures in SPSS: A Self-Help and Practical Guide to Preparing Theses, Dissertations, and Manuscripts*

Edited by: John R. Slate, Ana Rojas-LeBouef

URL: <http://cnx.org/content/col11292/1.6/>

License: <http://creativecommons.org/licenses/by/3.0/>

Module: "Introduction: Why a Book on Statistical Help for Graduate Students and Faculty?"

By: John R. Slate, Ana Rojas-LeBouef

URL: <http://cnx.org/content/m37280/1.2/>

Pages: 1-3

Copyright: John R. Slate, Ana Rojas-LeBouef

License: <http://creativecommons.org/licenses/by/3.0/>

Module: "Calculating Descriptive Statistics"

By: John R. Slate, Ana Rojas-LeBouef

URL: <http://cnx.org/content/m37276/1.9/>

Pages: 5-24

Copyright: John R. Slate, Ana Rojas-LeBouef

License: <http://creativecommons.org/licenses/by/3.0/>

Module: "Calculating a Nonparametric Pearson Chi-Square"

By: John R. Slate, Ana Rojas-LeBouef

URL: <http://cnx.org/content/m37277/1.5/>

Pages: 25-34

Copyright: John R. Slate, Ana Rojas-LeBouef

License: <http://creativecommons.org/licenses/by/3.0/>

Module: "Calculating Correlations: Parametric and Non Parametric"

By: John R. Slate, Ana Rojas-LeBouef

URL: <http://cnx.org/content/m37278/1.7/>

Pages: 35-53

Copyright: John R. Slate, Ana Rojas-LeBouef

License: <http://creativecommons.org/licenses/by/3.0/>

Module: "Conducting a Parametric Independent Samples t-test"

By: John R. Slate, Ana Rojas-LeBouef

URL: <http://cnx.org/content/m37279/1.5/>

Pages: 55-69

Copyright: John R. Slate, Ana Rojas-LeBouef

License: <http://creativecommons.org/licenses/by/3.0/>

Module: "Conducting a Parametric Dependent Samples t-test (Paired Samples t-test)"

By: John R. Slate, Ana Rojas-LeBouef

URL: <http://cnx.org/content/m37328/1.6/>

Pages: 71-83

Copyright: John R. Slate, Ana Rojas-LeBouef

License: <http://creativecommons.org/licenses/by/3.0/>

Module: "Conducting a Nonparametric Independent Samples t-test"

By: John R. Slate, Ana Rojas-LeBouef

URL: <http://cnx.org/content/m37330/1.6/>

Pages: 85-99

Copyright: John R. Slate, Ana Rojas-LeBouef

License: <http://creativecommons.org/licenses/by/3.0/>

Module: "Conducting a Nonparametric Paired Samples t-test"

By: John R. Slate, Ana Rojas-LeBouef

URL: <http://cnx.org/content/m37333/1.5/>

Pages: 101-111

Copyright: John R. Slate, Ana Rojas-LeBouef

License: <http://creativecommons.org/licenses/by/3.0/>

Module: "Conducting a Parametric One-Way Analysis of Variance"

By: John R. Slate, Ana Rojas-LeBouef

URL: <http://cnx.org/content/m37334/1.4/>

Pages: 113-129

Copyright: John R. Slate, Ana Rojas-LeBouef

License: <http://creativecommons.org/licenses/by/3.0/>

Module: "Conducting a Nonparametric One-Way Analysis of Variance"

By: John R. Slate, Ana Rojas-LeBouef

URL: <http://cnx.org/content/m37335/1.6/>

Pages: 131-147

Copyright: John R. Slate, Ana Rojas-LeBouef

License: <http://creativecommons.org/licenses/by/3.0/>

Module: "Standardized Skewness and Standardized Kurtosis Coefficient Calculator"

By: John R. Slate, Ana Rojas-LeBouef

URL: <http://cnx.org/content/m37347/1.3/>

Page: 149

Copyright: John R. Slate, Ana Rojas-LeBouef

License: <http://creativecommons.org/licenses/by/3.0/>

Module: "Resources: Calculating Basic Statistics in SPSS"

By: John R. Slate, Ana Rojas-LeBouef

URL: <http://cnx.org/content/m37336/1.3/>

Pages: 151-152

Copyright: John R. Slate, Ana Rojas-LeBouef

License: <http://creativecommons.org/licenses/by/3.0/>

Calculating Basic Statistical Procedures in SPSS: A Self-Help and Practical Guide to Preparing Theses, Dissertations, and Manuscripts

Calculating Basic Statistical Procedures in SPSS: A Self-Help and Practical Guide to Preparing Theses, Dissertations, and Manuscripts, is authored by John R. Slate and Ana Rojas-LeBouef from Sam Houston State University. This book is written to assist graduate students and faculty members, as well as undergraduate students, in their use of the Statistical Package of the Social Sciences-PC (SPSS-PC) versions 15-19. Specifically, we have generated a set of steps and screenshots to depict each important step in conducting basic statistical analyses. We believe that this book supplements existing statistical texts in which readers are informed about the statistical underpinnings of basic statistical procedures and in which definitions of terms are provided. Accordingly, other than providing a few basic definitions, we assume that dissertation chairs/thesis directors, students, and/or faculty will obtain their own definition of terms. We hope you find this set of steps and screenshots to be helpful as you use SPSS-PC in conducting basic statistical analyses.

About Connexions

Since 1999, Connexions has been pioneering a global system where anyone can create course materials and make them fully accessible and easily reusable free of charge. We are a Web-based authoring, teaching and learning environment open to anyone interested in education, including students, teachers, professors and lifelong learners. We connect ideas and facilitate educational communities.

Connexions's modular, interactive courses are in use worldwide by universities, community colleges, K-12 schools, distance learners, and lifelong learners. Connexions materials are in many languages, including English, Spanish, Chinese, Japanese, Italian, Vietnamese, French, Portuguese, and Thai. Connexions is part of an exciting new information distribution system that allows for **Print on Demand Books**. Connexions has partnered with innovative on-demand publisher QOOP to accelerate the delivery of printed course materials and textbooks into classrooms worldwide at lower prices than traditional academic publishers.